Identification of phytochemicals Rhinovirus

A THESIS SUBMITTED FOR THE PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF

BACHELOR OF TECHNOLOGY

IN

BIOTECHNOLOGY

By:

S MEGHASMITA

(Roll no. 111BT0588)

Under the guidance of:

Dr. SIRSENDU SEKHAR RAY

Department of Biotechnology and Medical Engineering

National Institute of Technology Rourkela

2015

NATIONAL INSTITUTE OF TECHNOLOGY

ROURKELA

DEPARTMENT OF BIOTECHNOLOGY AND MEDICAL ENGINEERING

CERTIFICATE

This is to certify that the thesis entitled "*Identification of phytochemicals for immunity against Rhinovirus caused Common cold"* submitted by Ms. S Meghasmita [Roll no. 111BT0588] in partial fulfillment of the requirements for the award of the degree of Bachelor of Technology in Biotechnology at National Institute of Technology, Rourkela is an authentic work carried out by her under my guidance.

To the best of my knowledge the matter embodied in the thesis has not been submitted to any other University/Institute for the award of any degree or diploma.

Date:

Dr. Sirsendu Sekhar Ray

Department of Biotechnology and Medical Engineering

National Institute of Technology Rourkela

ACKNOWLEDGEMENT

I avail this opportunity to extend my sincere appreciation and gratitude to my guide Dr Sirsendu Sekhar Ray, Biotechnology and Medical Engineering Department, for his invaluable academic and professional guidance, constant encouragement and kind help at different stages for the execution of this project. I want to thank Mr Praveen, Ms Priyanka Goyal and Mr Rohan V Bhagwat who guided me throughout the project. I want to thank my friends Anamika, Samapika and Smaranika who always supported me and stood by me everywhere. Last but not the least; I want to thank my family for their blessings.

S. Meghasmita

B.Tech (Biotechnology Engineering)

National Institute of Technology, Rourkela

CONTENTS

LIST OF TABLES AND FIGURES

ABSTRACT

Human rhinoviruses (HRVs), causative agents for over half of cold and cold like illnesses, also responsible for huge economic loss as it leads to loss in working hours. Our enhanced understanding of the complex genomic structure of the virus has been a direct result of advances in molecular methods and has resulted in characterization of three genetically distinct HRV groups, designated groups A, B, and C, within the genus Enterovirus and the family Picornaviridae. The HRV caused common cold is one of the most widespread viral infections for which no substantial treatment is available till now as there exists more than 100 serotypes of this virus. Research solutions so far have focused on targeting the conserved capsid proteins and the lipids for manufacture of viral membranes. I aim to target the viral RNA replication inside the host by finding a lead molecule obtained from any plant with any known antiviral property; which would bind to the Rdrp (RNA dependent RNA polymerase) site of the virus and hence stop RNA replication from taking place. To achieve this first a library of all the phytochemicals obtained from all the plants having antiviral properties is prepared. Using *in silico* approach, the phytochemicals are screened for toxicity and then the lead molecule having the highest affinity for the virus is found. Once the lead molecule it identified, Lead optimization is performed which can be used to manufacture a desired drug. Also some structural dynamics analysis are performed on the virus along with the lead molecule and also compared with the structural dynamics of other structural proteins of the body.

Keywords: Human Rhinovirus, phytochemicals, lead molecule, docking, structural dynamics

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1 INTRODUCTION

HRV, while once only thought to be the causative agent for relatively benign upper respiratory tract illness; it is now explained to have connections to exacerbations of chronic pulmonary disease, asthma development, and, more recently, fatal pneumonia in elderly and immune-compromised adults as well as severe bronchiolitis in infants and children. A recent advance in molecular methods has led to better understanding of the illness spectrum of HRVs which has led to the efficient characterization and detection of HRV strains and groups. Multiplex PCR-based assays for respiratory viruses detection that include HRVs are being adopted by quite an increasing number of clinical laboratories. Currently, there are no approved antiviral agents for HRV infection prevention or treatment. In natural setting drug toxicities, drug interactions, and a lack of efficacy have limited clinical trials of antiviral therapies when applied. There exists too much variability for over 100 serotypes at the antigenic sites, as a result of which the vaccine development has been hampered. Thus, the HRV infection treatment focuses mainly on symptoms relief and psychological consolation.

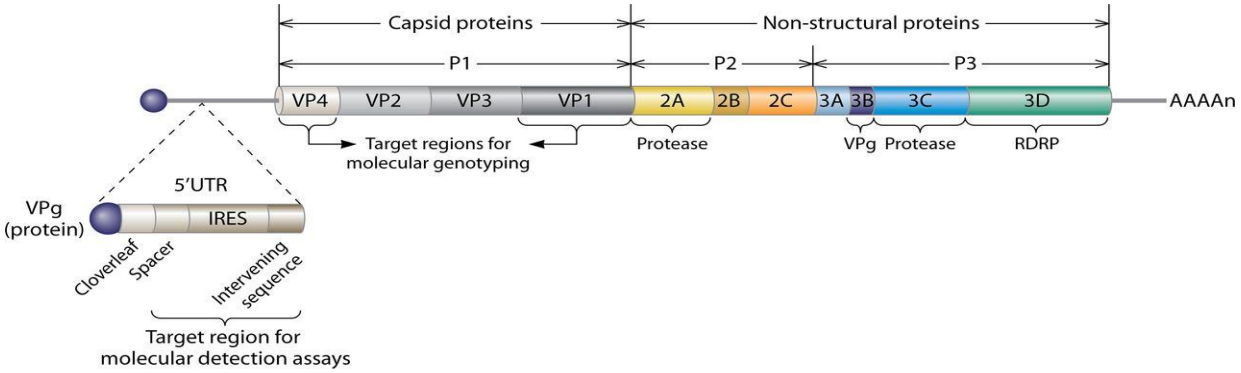


Fig 1.

1.1.1 Infection pathway for Rhinovirus:

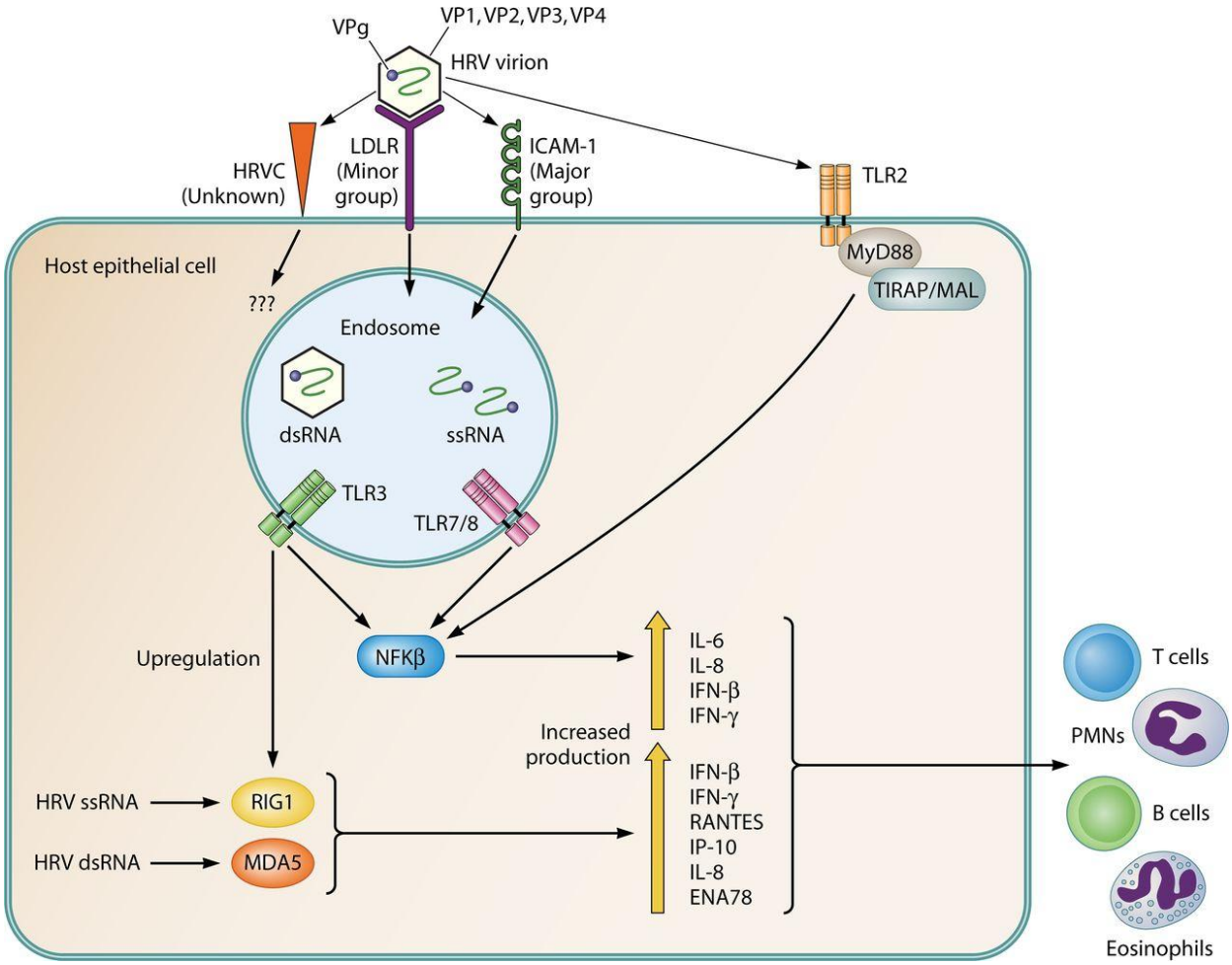


Fig. 2

HRVC- The third species of HRV, recently identified, whose pathogenesis is still largely unknown and yet to be discovered.

LDLR- The low density lipoprotein receptor (LDLR) gene family is made up of cell surface proteins which are involved in receptor-mediated specific ligands endocytosis.

ICAM 1- (Intercellular Adhesion Molecule 1) a cell surface glycoprotein encoded by this gene is typically expressed on immune system and endothelial cells. It binds to type CD11a / CD18, or CD11b / CD18 intigrins and is also serves as receptor to Rhinovirus.

TLR2- This gene codes a protein which is a member of the Toll-like receptor (TLR) family that plays the most important role in innate immune activation and pathogen recognition. Pathogen-associated molecular patterns (PAMPs) are recognized that are expressed on infectious agents, and the production of cytokines is mediated which is necessary for the effective immunity development. Most abundantly expressed in peripheral blood leukocytes, this gene mediates host response to yeast and Gram-positive bacteria via NF-kappaB simulation.

TLR3- Mostly expressed in pancreas and placenta, this receptor is limited to the dendritic subpopulation of the leukocytes.Viral infection associated dsRNA is recognized by it, and it also induces the activation of NF-kappaB and the production of type I interferons. This gene uses alternative polyadenylation sites for generation of different length transcripts.

TLR7- This gene is most notably expressed in placenta, spleen, and lungs, and on chromosome X it lies in close proximity to TLR8.

MyD88- myeloid differentiation primary response 88: This gene encodes a cytosolic adapter protein which plays the main role in the adaptive and innate immune response. This protein works as a necessary signal transducer in the Toll-like receptor and interleukin-1 signaling pathways. These pathways regulate that activation of a number of pro-inflammatory genes. The encoded protein consists of an N-terminal death domain and a C-terminal Toll-interleukin1 receptor domain.

Increased susceptibility to pyogenic bacterial infections is a symptom of patients with a defect in this gene.

TIRAP/MAL- The TIR adaptor protein is encoded by this gene which is involved in the TLR4 signaling pathway of the immune system. It activates MAPK1, MAPK3, NF-kappa-B and JNK, that then results in inflammatory response followed by cytokine secretion.

NFκβ- nuclear factor of kappa light polypeptide gene enhancer in B-cells: NFKB, activated by various extra and intra-cellular stimuli such as ultraviolet irradiation, cytokines, oxidant-free radicals, and viral or bacterial products is a transcription regulator. Activated NFKB stimulates the expression of genes involved in a wide variety of biological functions by translocating into the nucleus.

RIG1/MDA5 - RIG-I (retinoic-acid-inducible protein 1, also known as Ddx58) and MDA-5 (melanoma-differentiation-associated gene 5, also known as Ifih1 or Helicard) sense double-stranded RNA (dsRNA), a replication intermediate for RNA viruses, leading to production of type I interferons (IFNs) in infected cells. Viral dsRNA is also recognized by Toll-Like receptor 3 (TLR3) which is expressed on the cell surface membrane or endosomes.

Recognition of dsRNA by RIG-I/MDA-5 or TLR3 is cell-type dependent.

1.1.2 Infection mechanism

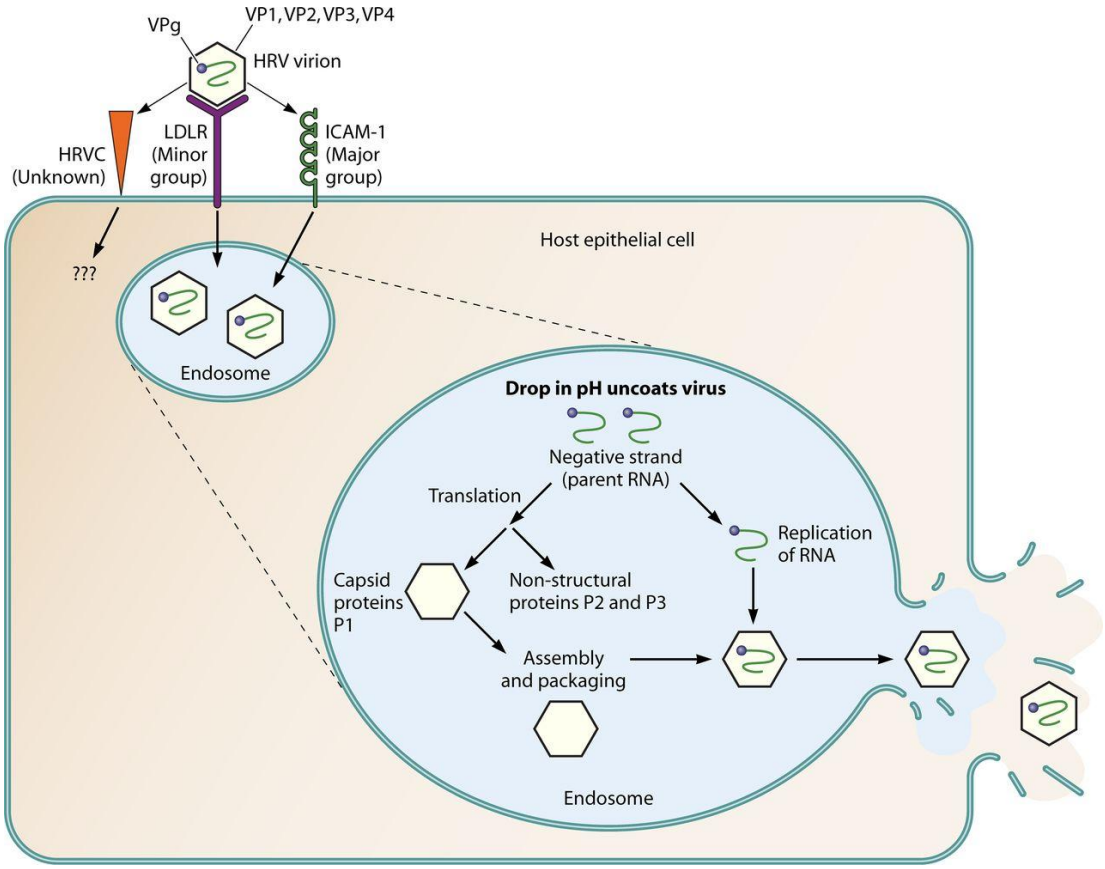


Fig 3.

Once the virus enters the human body and infects the endosome, the capsid coating of the virus is removed by the change in pH and the negative RNA strands are exposed. These ssRNA translate to give rise to capsid proteins for the manufacture of viral capsid. Simultaneously the ssRNA replicates to form more copies of the RNA.

The capsids thus manufactured and the replicated RNA molecules undergo assembly and packing to form new viruses exactly similar to the one that started the infection. Thus the reaction continues like this and more number of virus particles are produced.

1.2 LITERATURE REVIEW

1.2.1 Immunity against the conserved domains of the RV family: No common conserved regions between species A and B have been found but within both the species three highly conserved domains have been identified which include amino acids 1–191 and 243–297 in the N-terminus of the polyprotein, the C-terminal domain of the RNA polymerase and VP0. VP0 is the natural precursor of the two N- terminus regions which lie within the VP4 and VP2 capsid proteins. The scientists adopted a T helper1/T cytokine 1 oriented response against the viral infection caused by this VP0 for a better disease outcome. They suggest that the Th1 cell responses to HRV will prove a safe strategy for preventing HRV induced disease which awaits confirmation in a clinical trial.

1.2.2 The capsid proteins of all the serotypes of the HRV were analysed and the protein VP1 was found to contain highly conserved domains in all the serotypes. According to the same alignment, Individual ICAM-1 footprint amino acids, which earlier were suggested to play an important role in virus–receptor interactions, were found out not to be conserved.. Whereas, amino acid residues corresponding to the documented ICAM-1 footprint showed classification of the two receptor groups unequivocally. A number of similarity and divergence was found in the amino acid sequences found in the hydrophobic pocket of the VP1 capsid protein.

1.2.3 *Roulin et al* showed that rhinovirus replication depends on host factors driving phosphatidylinositol 4-phosphate (PI4P)-cholesterol counter-currents at viral replication membranes. Depending on the type of virus, replication required phosphatidylinositol 4-kinase class 3beta (PI4K3b), cholesterylesterasehormone-sensitive lipase (HSL) or oxysterol-binding protein (OSBP)-like 1, 2, 5, 9 or 11 associated with lipid droplets, endosomes or Golgi. Replication invariably required OSBP1, which shuttles cholesterol and PI4P between ER and Golgi at membrane contact sites. Infection also required ER-associated PI4P phosphatase Sac1 and phosphatidylinositol (PI) transfer protein beta (PITPb) shunting PI between ER-Golgi. These data support a PI4P-cholesterol counter-flux model for rhinovirus replication.They present evidence for a lipid counter-current model boosting rhinovirus infection. High concentrations of phosphatidyl-inositol (PI) 4-phosphate are built-up on Golgi membranes and used by oxysterol-binding protein-1, a sterol-dependent transducer in cytokine signaling to load cholesterol onto viral replication membranes at ER-Golgi contact sites using lipid droplet cholesteryl-esters.

1.3 The Rhinovirus structure used for the experiment:

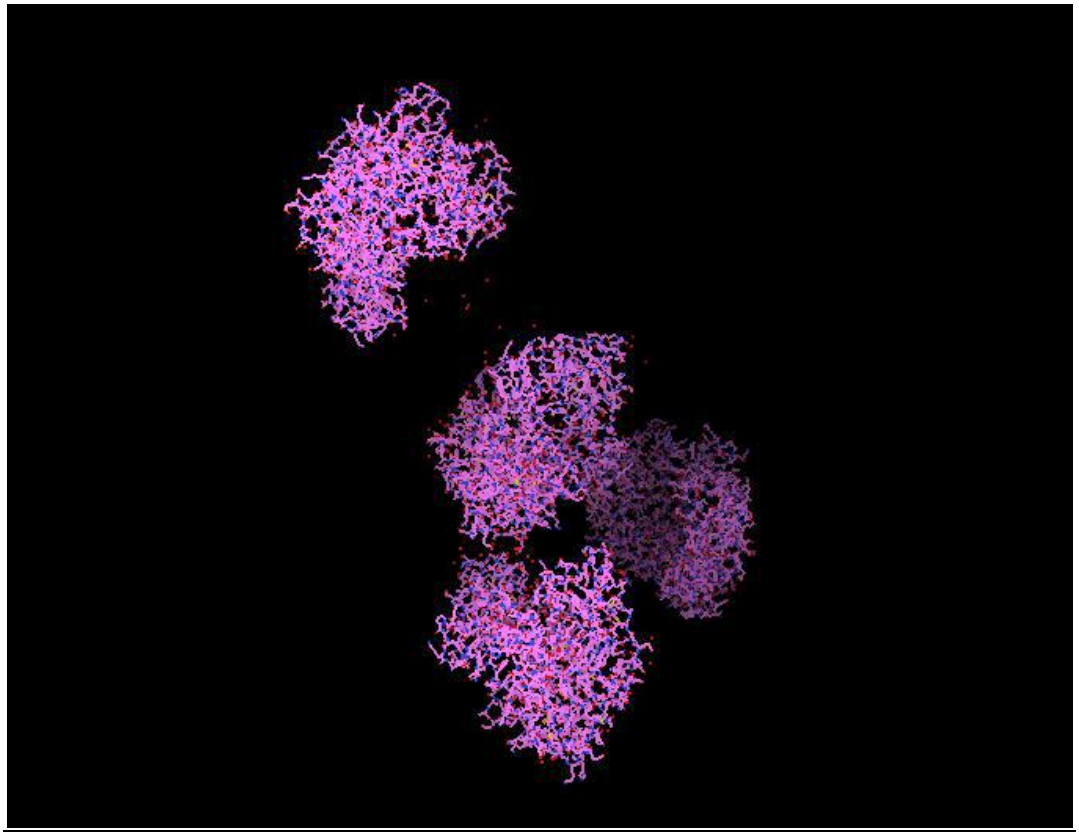


Fig 4.Rhinovirus 16 polymerase elongation complex (r1_form)

A vital role in the growth of RNA viruses is played by RNA dependent RNA polymerase where they are responsible for genome replication, but is done so with rather low fidelity which allows for the rapid adaptation to different host cell environments. These polymerases also act as a great target for antiviral drug development. However, a lack of detailed structural information about functional polymerase-RNA complexes and the structural changes that take place during the elongation cycle has hampered our understanding of fidelity determinants and also our drug discovery efforts.

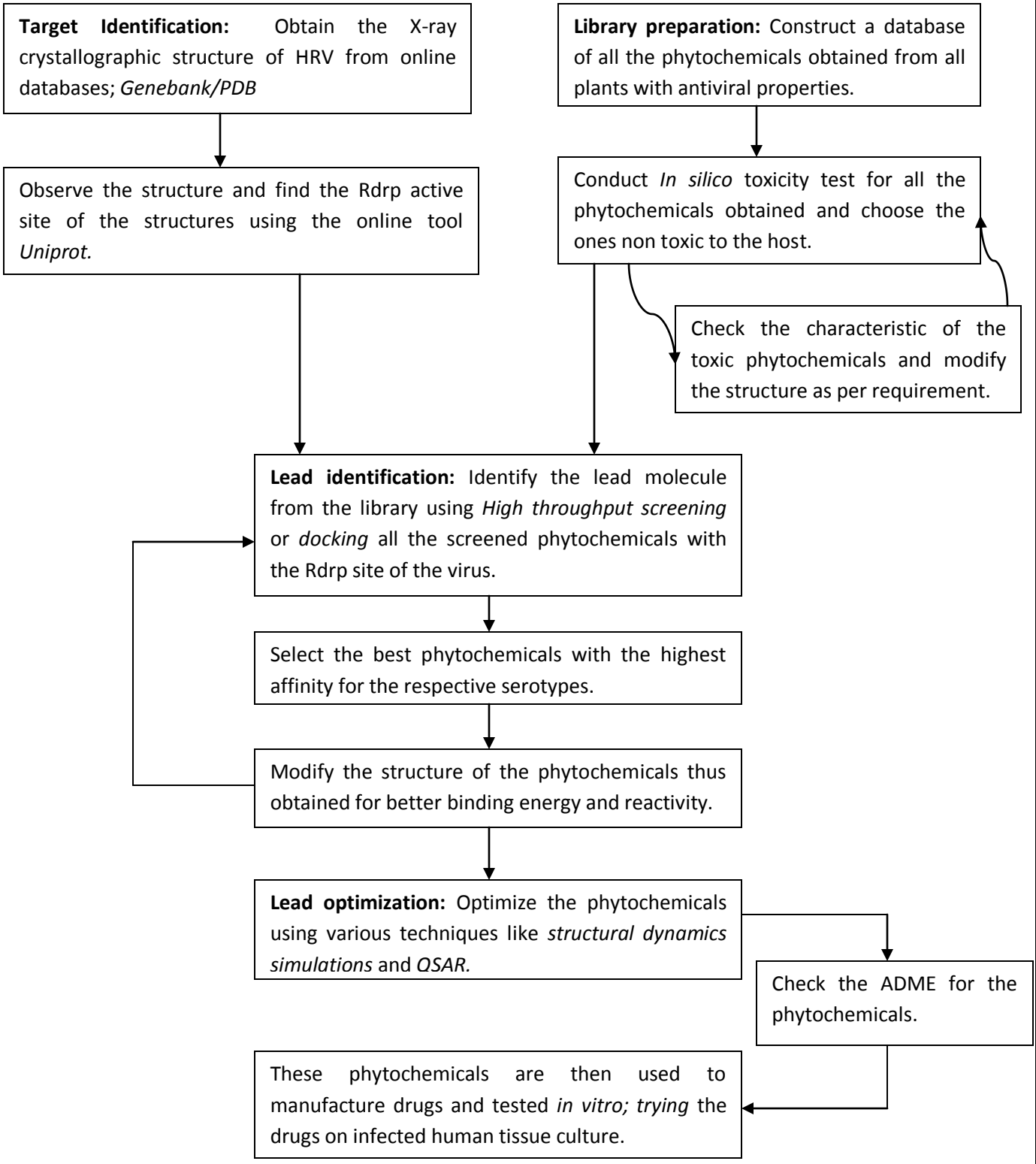
CHAPTER 2

OBJECTIVES AND WORK PLAN

2.1 OBJECTIVES:

- Preparation of a database of all the phytochemicals obtained from Indian plants having some or the other antiviral properties.
- Selection of the protein of interest (HRV) and performing docking with all the listed phytochemicals to identify the lead molecule.
- To perform lead optimization.

2.2 WORKPLAN:


Target Identification: Obtain the X-ray crystallographic structure of HRV from online databases; *Genebank/PDB*
Library preparation: Construct a database of all the phytochemicals obtained from all plants with antiviral properties.
Observe the structure and find the Rdrp active site of the structures using the online tool *Uniprot.*
Conduct *In silico* toxicity test for all the phytochemicals obtained and choose the ones non toxic to the host.
Check the characteristic of the toxic phytochemicals and modify the structure as per requirement.
Lead identification: Identify the lead molecule from the library using *High throughput screening* or *docking* all the screened phytochemicals with the Rdrp site of the virus.
Select the best phytochemicals with the highest affinity for the respective serotypes.
Modify the structure of the phytochemicals thus obtained for better binding energy and reactivity.
Lead optimization: Optimize the phytochemicals using various techniques like *structural dynamics simulations* and *QSAR.*
Check the ADME for the phytochemicals.
These phytochemicals are then used to manufacture drugs and tested *in vitro; trying* the drugs on infected human tissue culture.

CHAPTER 3

MATERIALS AND METHODS

3.1 Sofwares required:

1. **Argus lab**- for editing the structure of the protein or the ligand obtained from the Protein Data Bank or Uniprot. Similarly the ligand structure can be edited. Editing the structure such as, removing the water molecules or any other drugs (ligand) attached to the protein of choice etc.

2. **Chimera 1.8.1**- for viewing the protein or the ligand structure or the docked structure. The file should be in .pdb format. This is also used as an editor when the construction of the ligands are done.

3. **MGL tools 1.5.6**- for doing the docking of the chosen protein and ligand.

4. **Cygwin terminal**- used for docking analysis. It converts the grid parameter files (.gpf) and docking parameter files (.dpf) into readable formats such as the .glg and .dlg format. Thus getting the .pdb and .pdbqt file format of the docked molecule.

5. **Other websites** like Uniprot, PDB, PreADMET, PRODRG-server, PuBChem, KnapSack_3D, SMILES Translator (pdb to mol conversion) etc.

6. For normal mode analysis- http://lorentz.immstr.pasteur.fr/index1.php

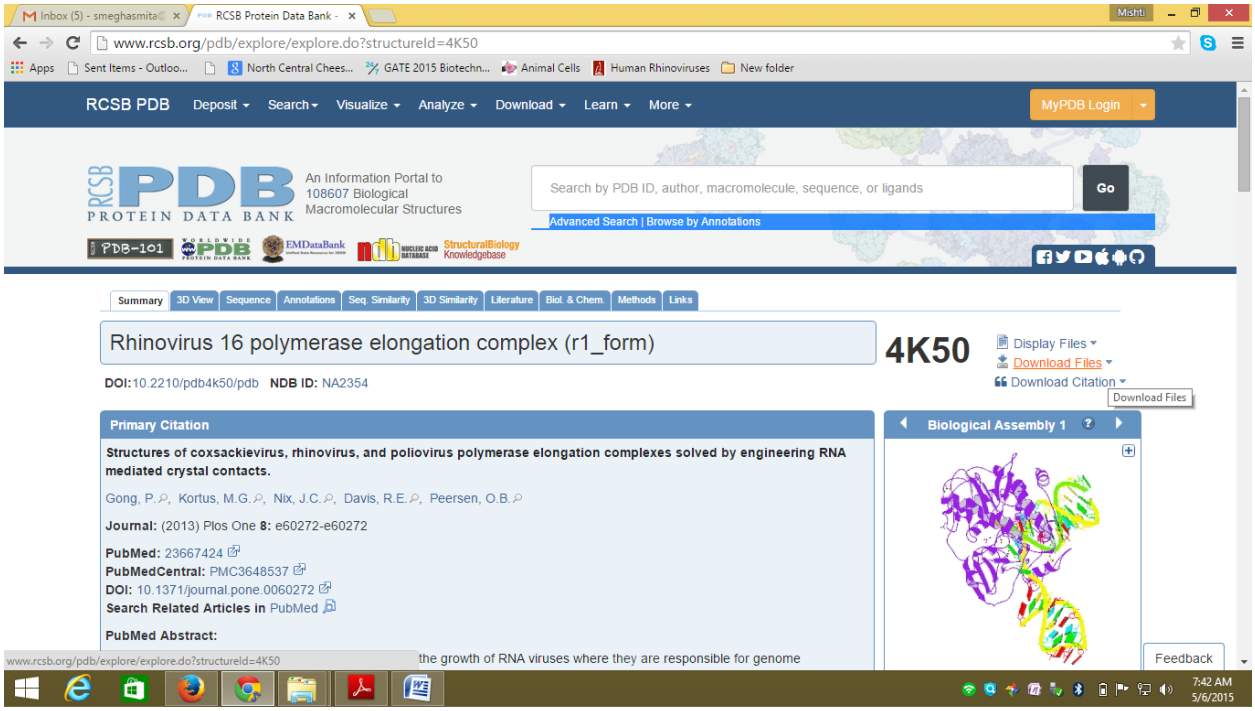
7. For elastic mode analysis- http://enm.lobos.nih.gov/

8. For Gaussian network model analysis- http://ignm.ccbb.pitt.edu/GNM_Online_Calculation.htm

9. For flexibility analysis - http://mmb2.pcb.ub.es/FlexServ/viewData.php?sessId=8F1A2A1E0E2F5C7B9A7E

3.2 Methodology:

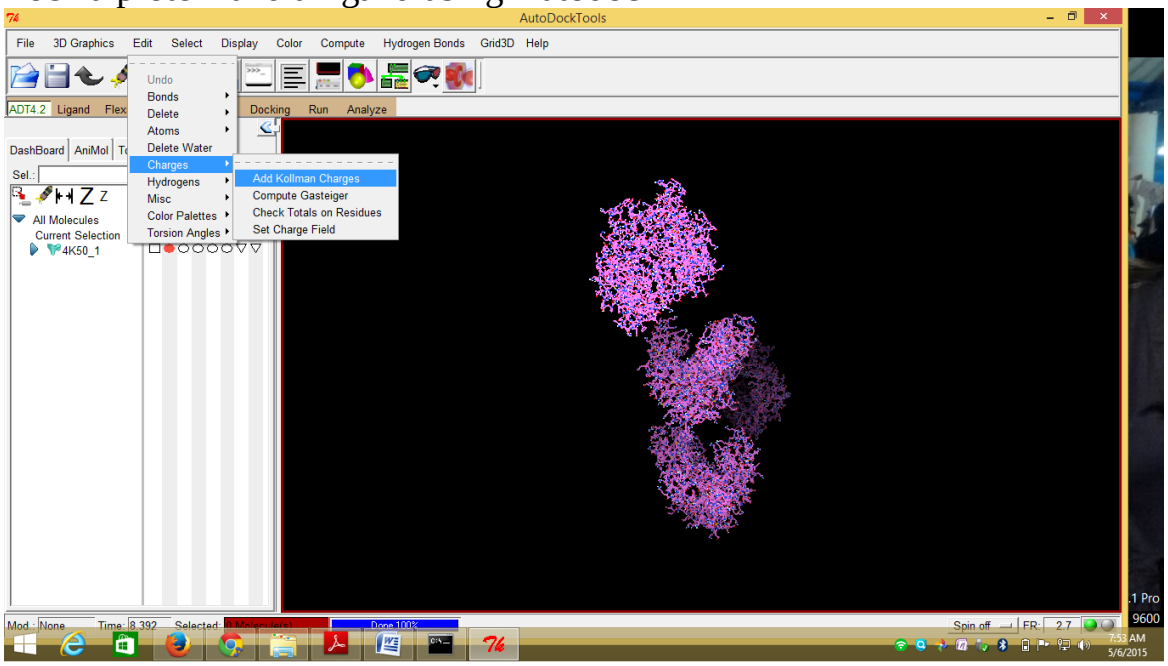
- Preparation of a database of all the Indian plants with some antiviral property, listing all its phytochemicals
- **3.2.1 Docking**
- Get the protein for the docking study: The protein of interest is searched in Uniprot or Protein Data Base (PDB) and thus the required protein structure is downloaded in the pdb file (text) format. If suppose we get the structure of our protein from the Uniprot or PDB, such that the ligand (like drugs etc.) is already attached to it then we can remove such ligands from our protein of interest by using the ArgusLab.exe software.



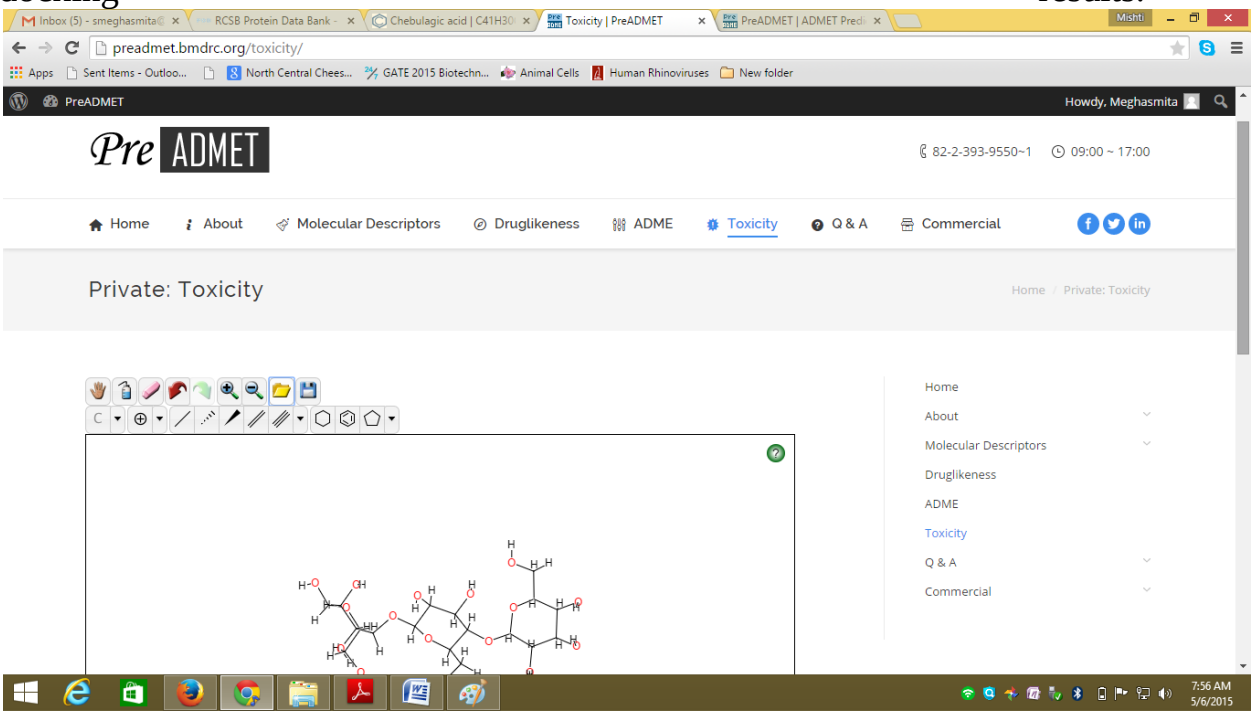
- Also, all the water molecules are deleted so that the required protein structure has all the hydrophilic sites free in it. Now the protein is ready to be docked.

- Get the ligand for the docking study: Ligands are generally searched in different ligand search engines such as PubChem, KnapSack_3D and many others available. After the desired ligand is selected it is downloaded and thus saved in .pdb file format. To save the ligands in .pdb file format the PRODRG-server is used. Thus now the ligand is also ready for docking.

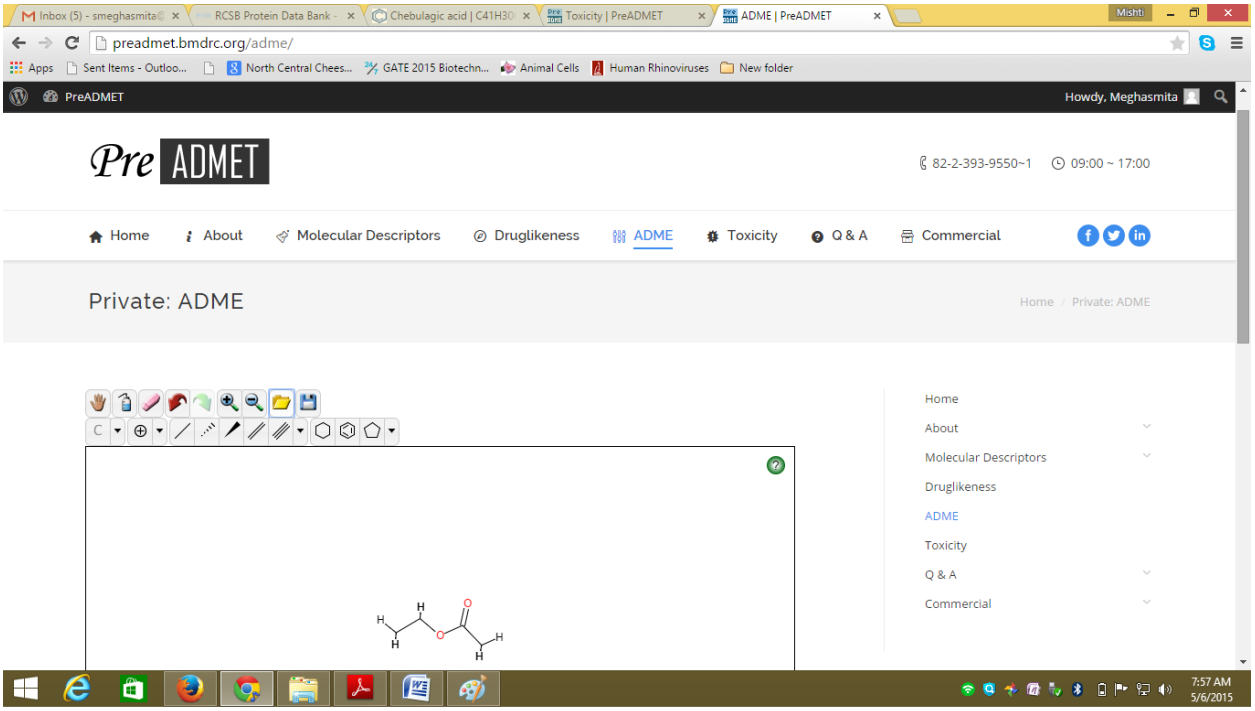
- Dock a protein and a ligand using Autodock



- **3.2.2 Toxicity test**
- The toxicity of the molecules is then checked for the molecules with best docking results.



- **3.2.3 ADME test:** Following toxicity, ADME is predicted for those molecules



- **3.2.4 Structural dynamic analysis**
- After the best molecule is obtained, the structural dynamics of the Rhinovirus bound to the best molecules is analyzed by using a number of online tools like EMN, MNA and GNM calculators and compared to some common structural proteins.
- Also, the flexibility of the protein is analyzed using online tool flexserve.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 DOCKING RESULTS:

We performed docking of a large number of molecules from the database prepared. We didn't get very favorable results from the docking analysis. Here we list the 10 best docking results.

Phytochemical	Source	B.E (kcal/mol)	Ki(uM)
Chebulagic acid	Terminalia chebula	-8.81	3.5
Strictinin	Tea	-7.55	2.92
D-glucoside	A.vasica	-6.69	12.46
Pinocembrin	Honey	-6.67	12.85
Quercitrin	F.tataricium	-6.53	16.26
Theaflavin	Tea	-6.52	16.56
Myricitrin	R. paviflora	-6.3	24.28
Valine	A.vasica	-6.26	25.98
Isovitexin	A.vasica	-6.19	29.21
α-amyrin	A.vasica	-6.13	31.92

Table 1.

1. Of all the docks performed till now, Chebuagic acid is the molecule with the best results so far.
2. Chebulagic acid, extracted from the plant Terminalia chebula, exhibits greater than 50% inhibition of HBV DNA polymerase activity *in vitro* and has showed potential antiviral activity effects against duck hepatitis B virus. *T chebula* fruits show four immunodeficiency virus GA (I), type 1 (HIV-1) integrase inhibitors and three galloyl glucoses (II-IV). It also shows protection of epithelial cells against influenza A virus. It also demonstrated therapeutic activity against herpes simplex virus both *in vitro* and *in vivo* tests. Tannins from*T. chebula* are effective against potato virus x. (Anwesa Bag et al., 2013).

3. All the 10 molecules listed in the table have been proved to show some anti viral property against come or the other virus.

4. All these listed phytochemicals are targeted to inhibit the RNA replication and thus stop the rhinovirus from following the entire pathway.

4.2 TOXICITY RESULTS:

Compunds	algae_at	Ames test	Carcino_Mouse	Carcino_Rat	daphnia_at	hERG_inhibition
Chebulagic acid	1.45E-06	non-mutagen	negative	negative	0.0197525	ambiguous
Strictinin	0.00094343	non-mutagen	negative	negative	0.0801163	ambiguous
D-glucoside	0.241751	mutagen	negative	negative	31.0365	low risk
Pinocembrin	0.0617865	non-mutagen	negative	negative	0.144608	medium risk
Quercitrin	0.0203185	non-mutagen	negative	negative	0.55858	high-risk
Theaflavin	0.00157179	non-mutagen	negative	negative	0.108776	high-risk
Myricitrin	0.0143948	non-mutagen	negative	negative	0.632419	ambiguous
Valine	0.128195	mutagen	positive	positive	2.50339	low risk
Isovitexin	0.0252359	mutagen	negative	negative	0.79706	high-risk
Alpha-amyrin	0.00143161	non-mutagen	positive	positive	0.00742832	low risk

Table 2.

Algae_at is the *in vitro* toxicity test performed on algae. It shows the level of toxic accumulation in algae if it is treated with the chemical for 2 days.

Ames_test is a simple method for testing mutagenicity of a compound, which was suggested by Dr. Ames. It makes use of several strains of the bacterium *Salmonella typhimurium* that carry mutations in genes which are involved in histidine synthesis, so that histidine is required by them for growth. The variable being tested shows the mutagen's ability to cause a reversion to growth on a histidine-free medium.

Carcino_Mouse: 2 years carcinogenicity bioassay in mouse.

Carcino_Rat: 2 years carcinogenicity bioassay in rat.

Daphnia_at: Acute daphnia(fish) toxicity.

hERG_inhibition: in vitro Human ether-a-go-go related gene channel inhibition. The protein product of the human ether-a-go-go gene (hERG) is a potassium channel which when inhibited by some drugs may lead to cardiac arrhythmia. Inhibition of the cardiac human ether-a-go-go-related gene channel is a major safety liability in clinical practice as it is a problematic off-target pharmacological activity.

Analysis of the toxicity:

- Chebulagic acid, which is our best result from docking studies, shows good results for the toxicity test. It shows the lowest levels of toxicity in the algae test. Ames test shows it to be a non mutagen and its non toxic to both mouse and rat tests. Also, it shows very low toxic accumulation in daphnia. But, it shows ambiguous results for hERG inhibition, which means it may or may not present a risk to the humans. There are chances, if administered into the human body it might inhibit the potassium channel and lead to cardiac arrhythmia.
- Of all the molecules, D-glucoside, valine and alpha-amyrin show the lowest risk to cause human cardiac arrhythmia but D-glucoside and valine are found to be mutagens in the Ames test. Also, valine and alpha-amyrin are found to be carcinogens in the mouse and rat bioassays.
- Only one molecule with medium risk to hERG inhibition is found to be Pinocembrin. It shows good results for all the tests. Also, the Ames test shows it to be a non-mutagen. Thus, Pinocembrin is the best molecule found in the toxicity test.

- Chebulagic acid, strictinin and myricitrin show extremely low levels of toxicity in all the tests but are ambiguous when it comes to the hERG inhibition test.

4.3 ADME RESULTS:

Molecules	BBB	Caco2	HIA	MDCK	Plasma_Protein_Binding	Skin_Permeability
Chebulagic acid	0.066303 1	17.696	0	0.043415 5	100	-2.92384
Strictinin	0.029520 5	12.8541	0.822184	0.099712 3	100	-4.32686
D-glucoside	0.059386	2.56748	22.3550 5	0.655125	7.312029	-5.30926
Pinocembrin	0.907503	2.47438	92.3541 4	43.0464	98.454853	-3.40499
Quercitrin	0.036954	7.37247	24.9475 2	1.81906	64.952873	-4.64502
Theaflavin	0.035603	14.8721	26.2274 1	0.044782 8	100	-4.48468
Myricitrin	0.033428	6.14008	11.6464	1.52579	65.37898	-4.71467
Valine	0.410878	20.0466	75.2550 8	7.36034	88.619012	-3.41617
Alpha-amyrin	20.9638	47.1744	100	1.43045	100	-2.10144

Table 3

AD prediction interpretation

There are several methods used in drug selection process to check for intestinal absorbtion. Caco-2 model and MDCK cell are referred to as *in vitro* model for prediction of drug absorbtion, where the drug has been administered orally. PreADMET predicts the permeabilty of Caco-2 cell, MDCK cell, blood brain barrier, Human intestianl absorption, skin permeability and plasma protein binding.

Caco-2 cells are derived from human colon adenocarcinoma which has multiple drug transport pathways.

Caco-2 permeability in PreADMET is classified into three classes.

1. Permeability<4, then the molecule is having low permeability.
2. 4<permeability<70, then the molecule is having medium permeability.

3. Permeability>70, then the molecule is having higher permeability.

MDCK cell permeability

MDCK cell refers to Madin-Darby canine kidney cell. Since MDCK cells life span is less than the life span of Caco-2 cells, correlation between them is said to be high. MDCK cell system can be used as a good tool for rapid permeability screening.

MDK cell level of permeability in PreADMET can be classified into three classes,

1. permebiality<25= molecule has low permeability.
2. 25<Permebiality<500= molecule has medium permeability.
3. Permebiality>500= molecule has higher permeability.

Human intestinal absorption

The intestinal absorption is very important to find the potential candidate and PreADMET finds the absorption in percentage. Poorly absorbed compounds return 0 to 20%, moderately absorbed compounds return 20- 70% and well absorbed compounds return 70-100%.

Skin permeability

Skin permeability factor is very important in case of cosmetics for transdermal delivery of drugs. PreADMET predicts the *in vitro* skin permeability and the result is given as logKp.

Blood Brain Barrier

It predicts whether compounds passes across the blood-brain barrier or not as it is crucial in the pharmaceutical sphere because CNS-active compounds are the only substances which must cross the barrier. In PreADMET, one can predict rates for BBB penetration in vivo data.

PreADMET explains about the absorbtion to CNS as follows.

Classification	BB	logBB
High absorption to CNS	More than 2.0	More than 0.3
Middle absorption to CNS	2.0 to 0.1	0.3 to -1.0
Low absorption to CNS	Less than 0.1	Less than -1.0

Table 4: Absorption level in central nervous system

Plasma protein binding of a drug gives information on the drug action, deposition and efficacy.

In PreADMET, plasma protein binding capacity is referred to as

Classification	Plasma protein binding
Strongly bound chemicals	More than 90%
Weakly bound chemicals	Less than 90%

Table 5

ADME RESULT ANALYSIS:

- The best molecule of our docking result i.e., chebulagic acid, shows medium absorption to the human colon. It has low absorption in the kidney thus, can be easily excreted from the body. But, this molecule shows no absorption in the intestine which is not favorable for a drug. It shows very low permeability through the skin and strongly bound to the plasma but it also shows very low absorption to the CNS.

- Our best result from the toxicity test i.e., Pinocembrin shows medium absorption to the CNS and low absorption to the colon. It is very well absorbed by the intestines and shows medium absorption to the kidney. It is strongly bound to the plasma protein and has very low permeability to the skin. Thus, overall it shows very good result in the ADME test.

LEAD OPTIMIZATION: PROTEIN STRUCTURAL DYNAMICS RESULTS:

Proteins usually are dynamic molecules and keep on changing with time and residue. They have inherent flexibility which allows them to modify as required and interact with a number of molecules. Proteins owe their various conformational changes and response towards different molecules, to their dynamic nature.

The protein structural analysis is done using a number of online tools which give the Normal mode analysis, the elastic network mode analysis, Gaussian normal mode analysis and flexibility of the protein along each residue.

The normal mode analysis shows the change in the residue at every node. The entire protein is analyzed for a specific time period and nodes are specific time instances considered at which the analysis is done for each residue of the protein. Resultantly we get the conformational changes in protein at specific point of time.

The elastic network model is the harmonic approximation of free surface energy of the protein near to its stable conformation. It describes the low frequency collective motions in the protein.

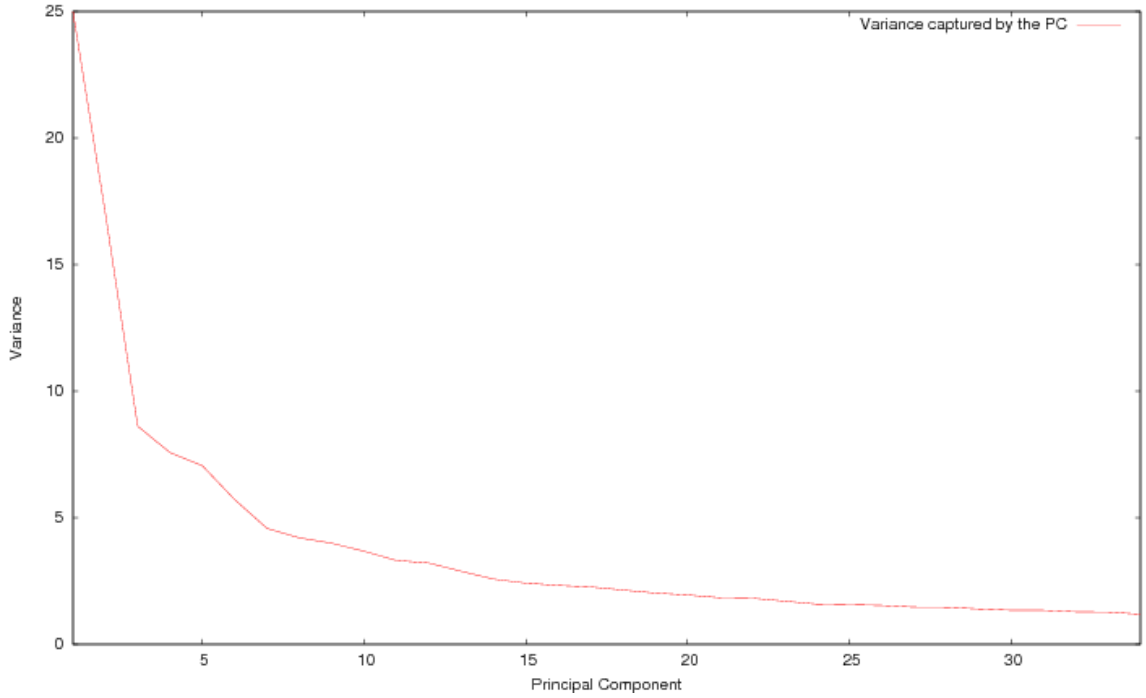
CC plots obtained from the Gaussian network model gives the cross correlations among the fluctuations in the residues. The values vary from +1 to -1.

+1 : red coloured. Shows that the fluctuation vectors are in the same direction also in the same sense.

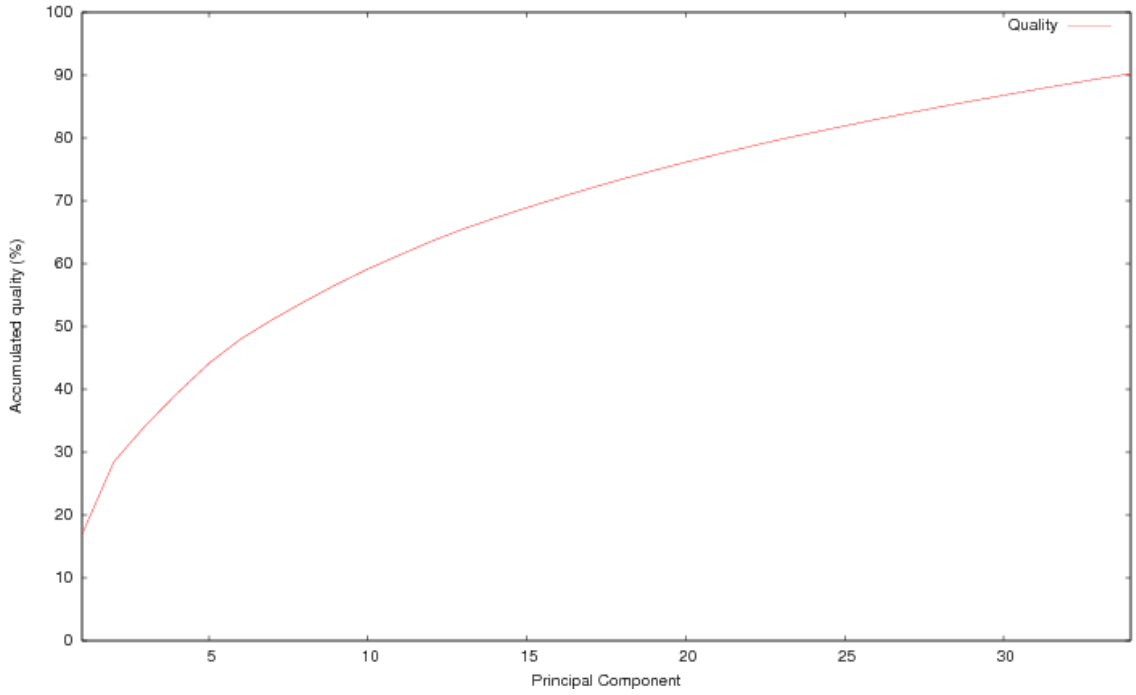
-1 : dark blue coloured. Shows that the fluctuation vectors are in the same direction but in the opposite sense.

- Here we first compute the flexibility of the HRV protein used to perform the experimentation.
- We then perform the structural dynamics simulation of the rhinovirus and compare it with four structural proteins i.e. collagen 1, collagen 9, keratin and silk fibroin.
- By comparing we can find the structural similarity of the rhinovirus with the structural proteins in the body and thus can find the physical strength of the rhinovirus.
- We then carry out the structural dynamic simulations of the rhinovirus bound to the 10 best ligands obtained earlier and analyze what effect these ligands have on the physical properties of the protein.
- We also compare the ligand bound protein to the structural proteins to find out if there is any similarity exists because when treating the virus with the ligands we have to make sure that the structural proteins of our body are not affected.

Variance profile:



Dimensionality/Quality profile

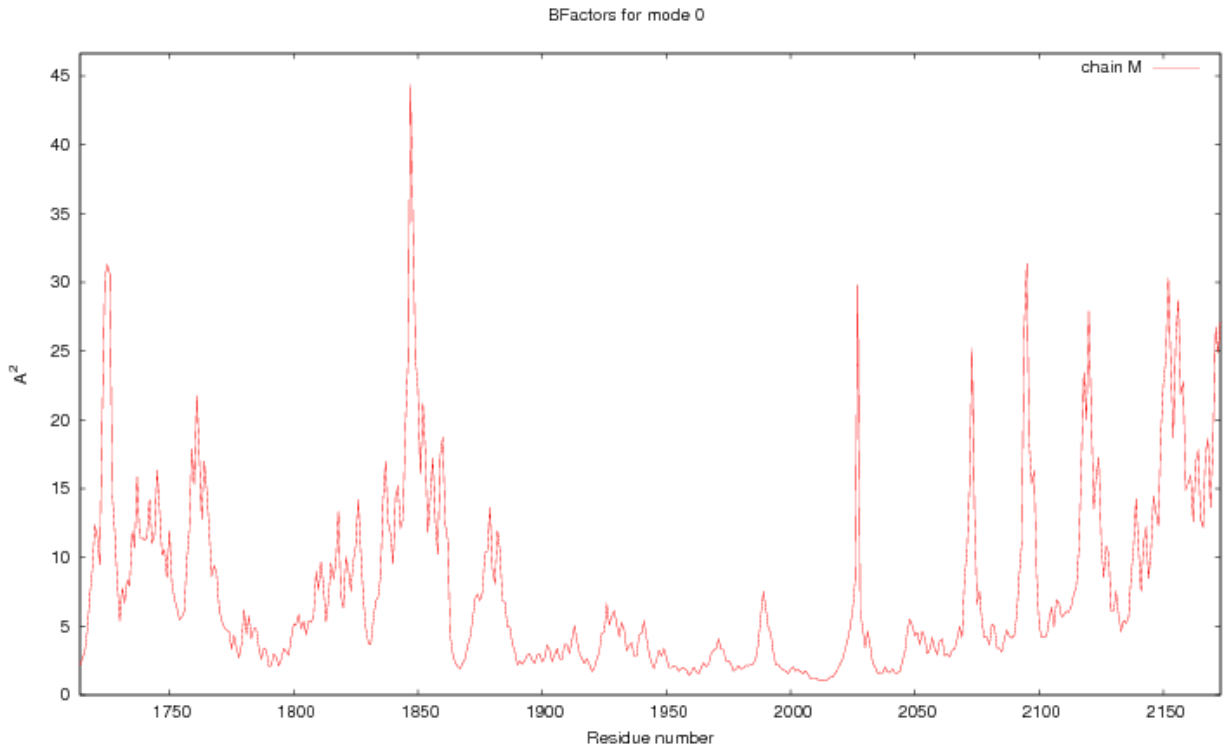


Size of essential space	38 eigenvalues
Total variance captured	147.770 Å²
Percentage of variance captured	90.24%

The first graph shows the total variance of the trajectory and how it is distributed along the different eigenvectors. This graph indicates the size of the flexibility space (higher the variance higher the flexibility) and how it is distributed in different modes. Clearly, Rhinovirus's flexibility decreases with the higher nodes.

The second graphic shows the percentage of explained variance for a given number of eigenvectors (quality) and the dimensionality of the sampled space. This graph indicates the complexity of the flexibility space, i.e. how many modes are necessary to explain the entire flexibility of the protein. Note that both the graphs provide physically-different information and that proteins might display a very complex pattern of flexibility (leading to large dimensionality) and at the same time be quite rigid (low variance), or have a large variance which can be fully explained by a very small number of modes. For HRV, the number of nodes required to explain the flexibility increases as the number of residues increases.

B factor profile:



B factor determines the flexibility of the molecule.

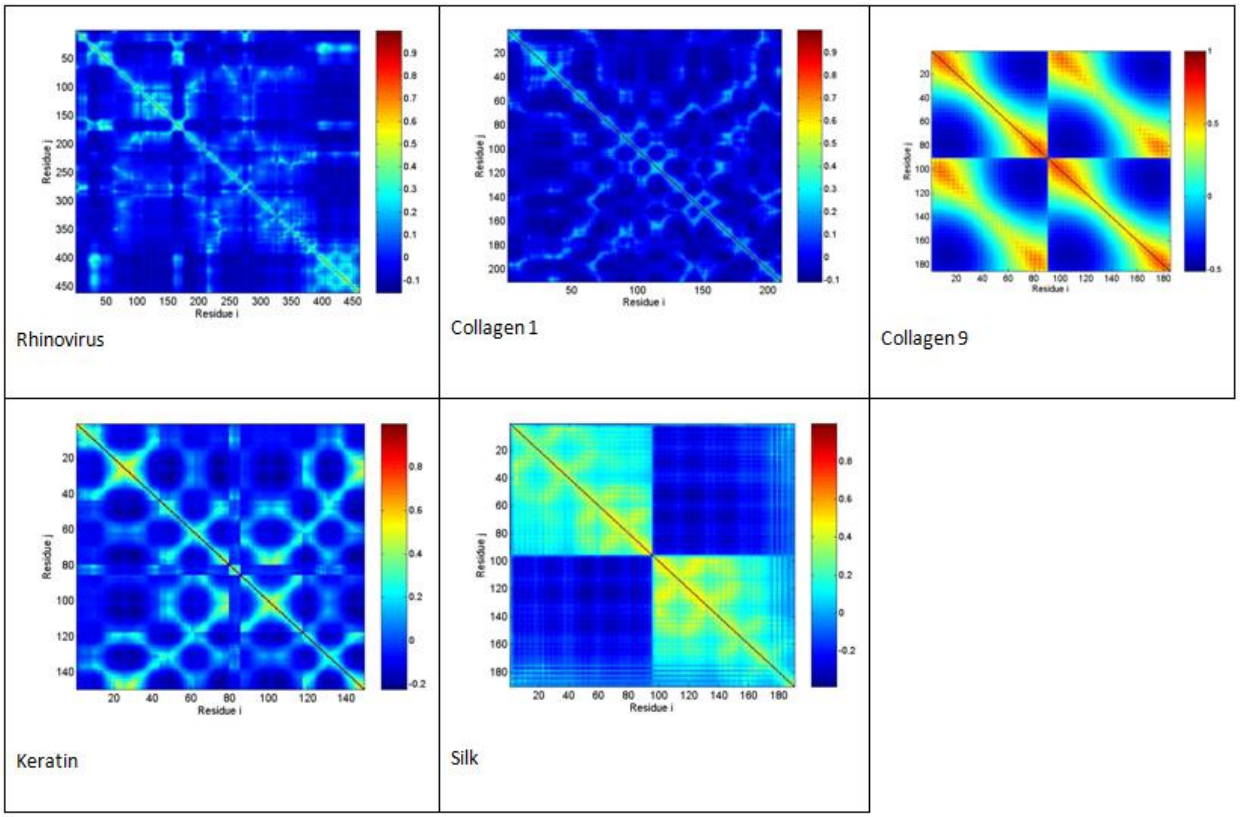
$$B - factor = \frac{8}{3}\pi^2 \langle \Delta r^2 \rangle$$

$\langle \Delta r^2 \rangle$ = the oscillations of residues around equilibrium positions.

B-factor profiles represent the distribution of residue harmonic oscillations. Very large-B-factors indicate very flexible residues that might display conformational changes along the trajectory, which is difficult to follow within the harmonic approximation implicit to B-factor analysis.

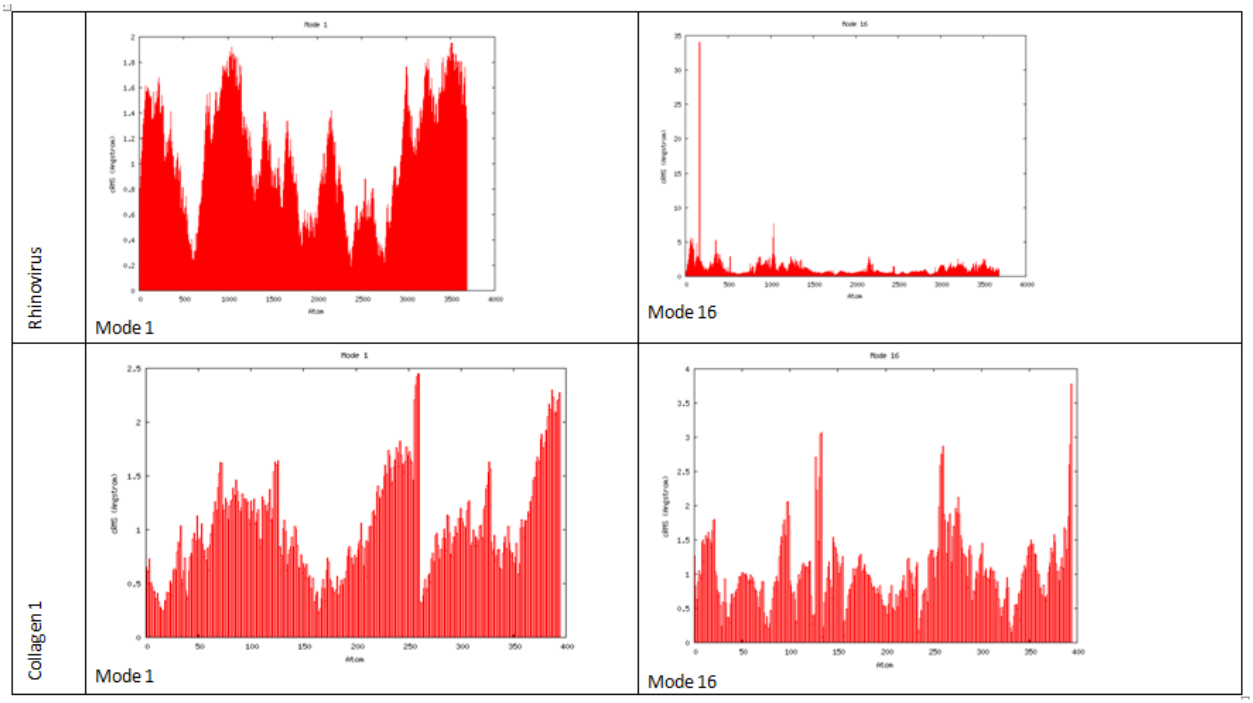
Our protein shows a variety of flexibility along each residue. From this we can infer that our protein is a very unstable one. Some of the residues show high flexibility while some show very low flexibility. It indicates that the protein is extremely dynamic in nature and keeps on changing its overall conformation constantly.

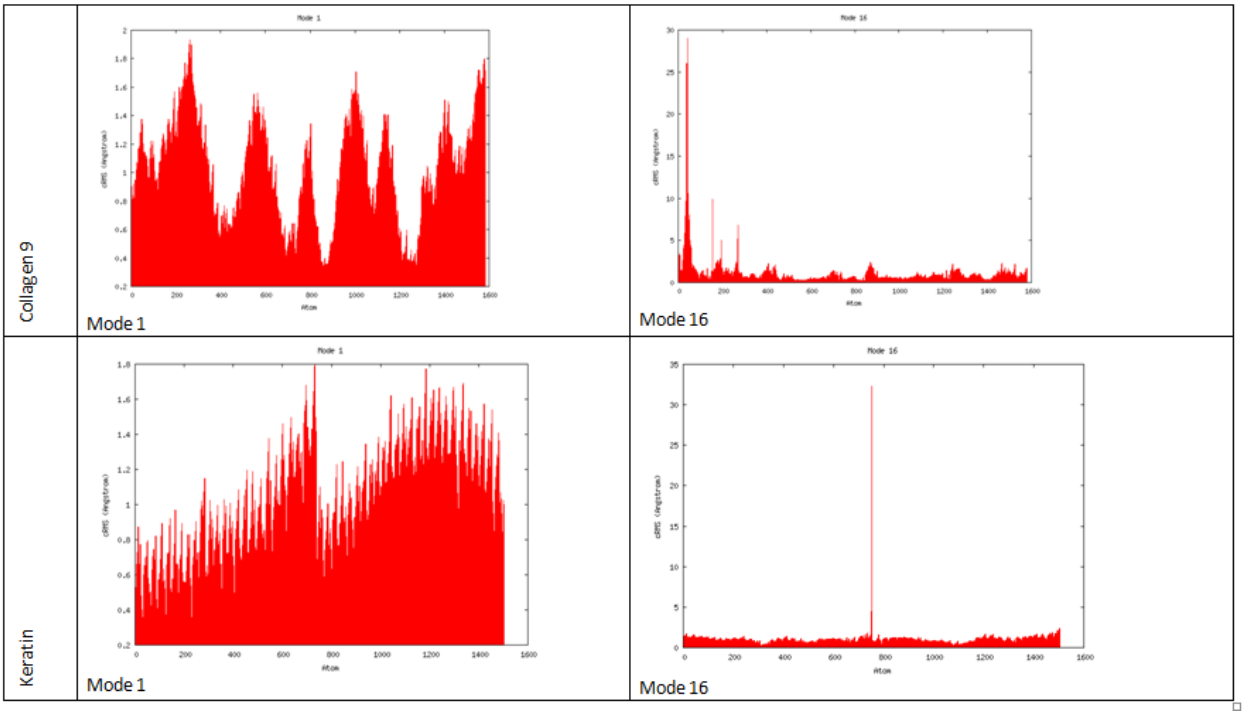
Comparing Rhinovirus with the structural proteins:



The above graphs are a result of the Gaussian network modeling and they depict that our viral protein shows very similar residue dynamics with collagen1 molecules. It also shows a little similarity with keratin. But the virus protein is found to be extremely different from collagen 9 and silk molecules.

In Rhinovirus, collagen 1 and keratin, majority of the residues show fluctuations in the same direction but in opposite sense.





Comparing the Normal mode analysis of Rhinovirus with the selected structural protein showed very different result from the one obtained from the Gaussian network model.

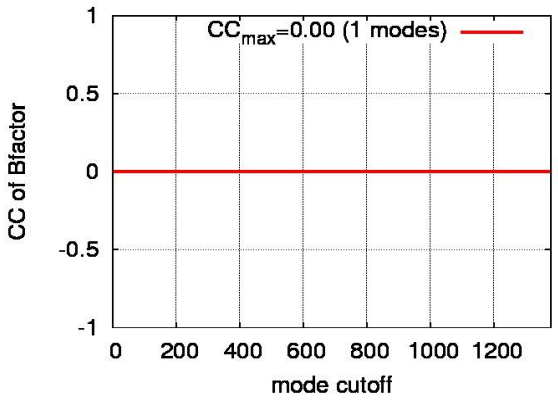
According to the Normal Mode Analysis (NMA), the viral protein shows dynamics very similar to those shown by the collagen 9 molecule. It shows striking difference from collagen 1 molecule.

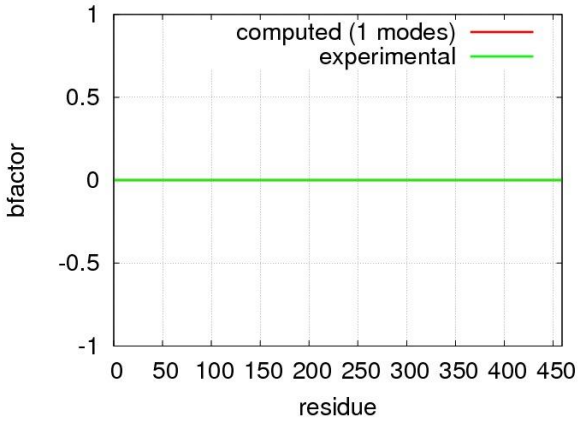
However, continuing from the Gaussian network model, it shows very dissimilar results from the other two structural proteins.

We then analyze the NMA, Elastic Network Model (ENM) and Gaussian network model for the Rhinovirus molecule bound to our top ligands obtained from the docking, toxicity and ADME test results.

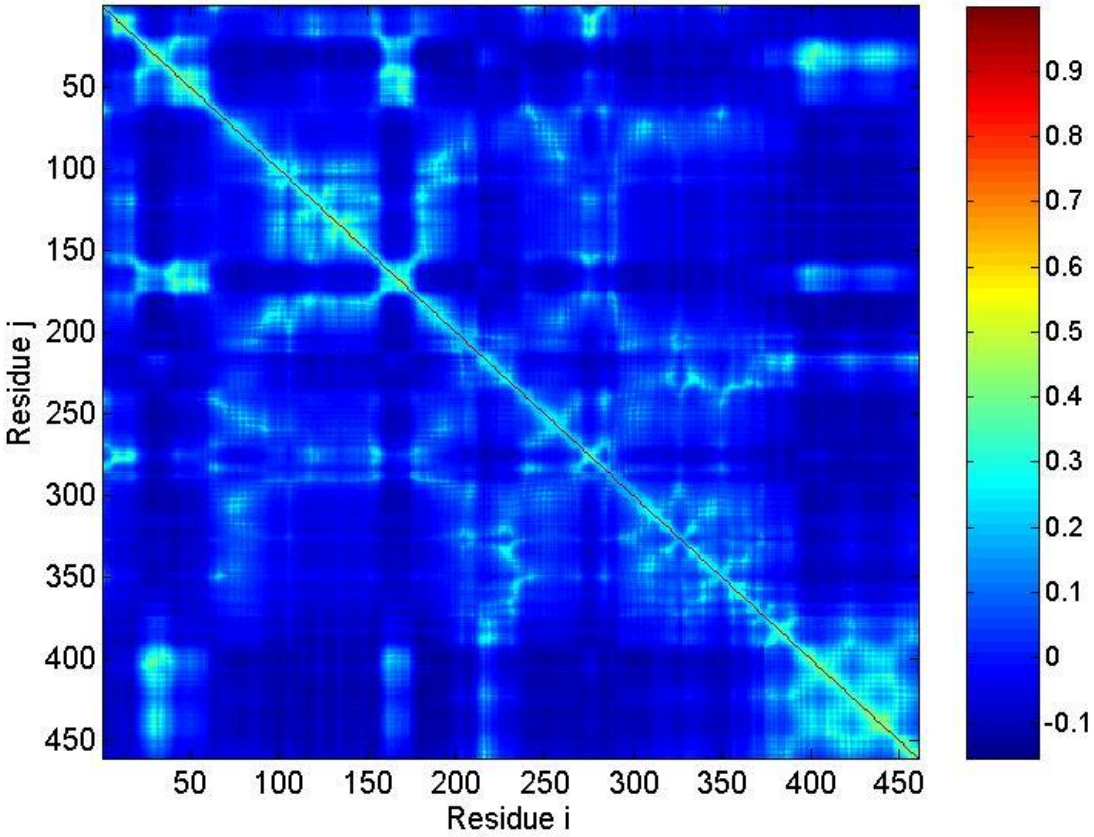
Chebulagic acid- Rhinovirus complex:

Elastic Normal mode:

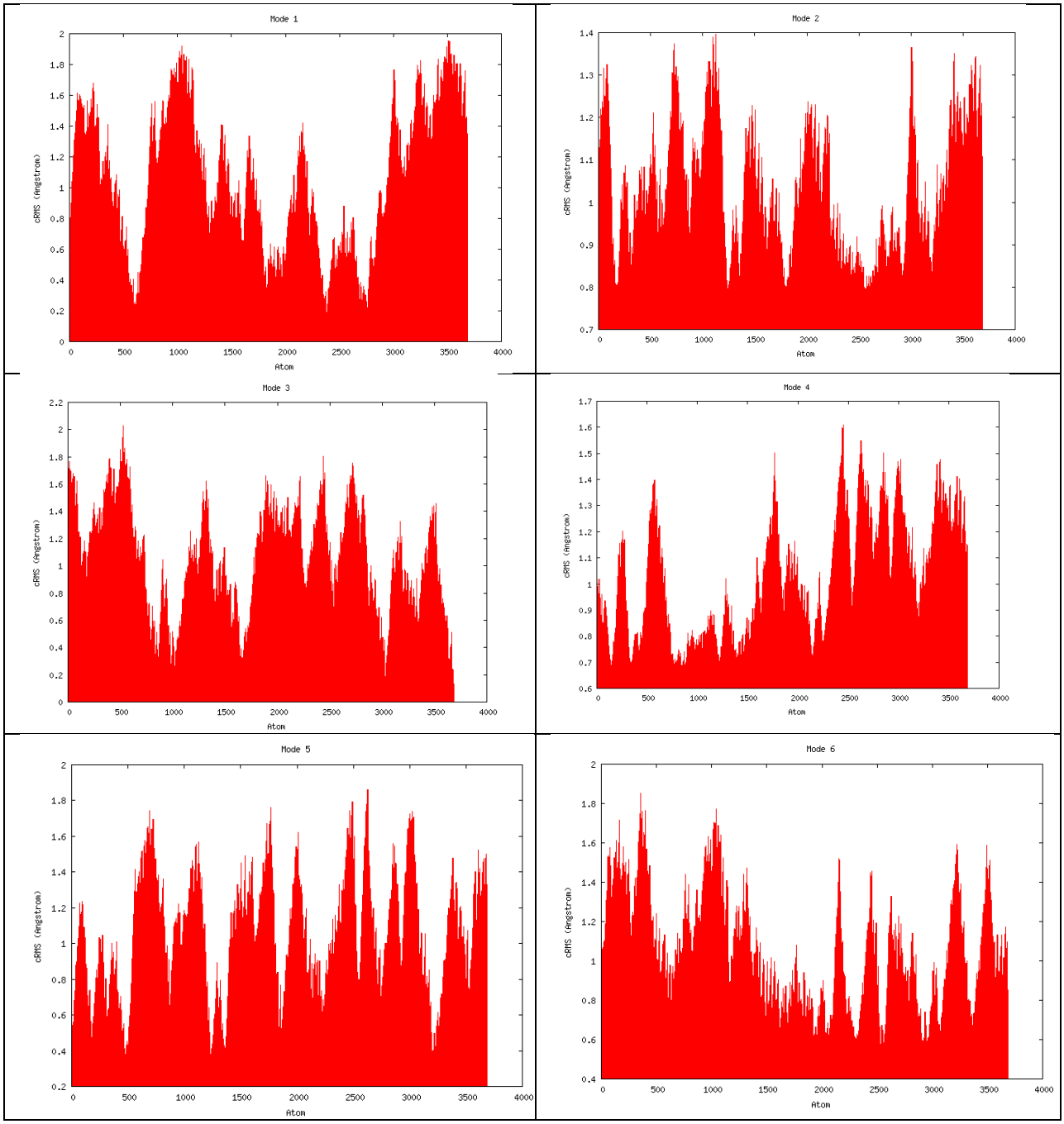


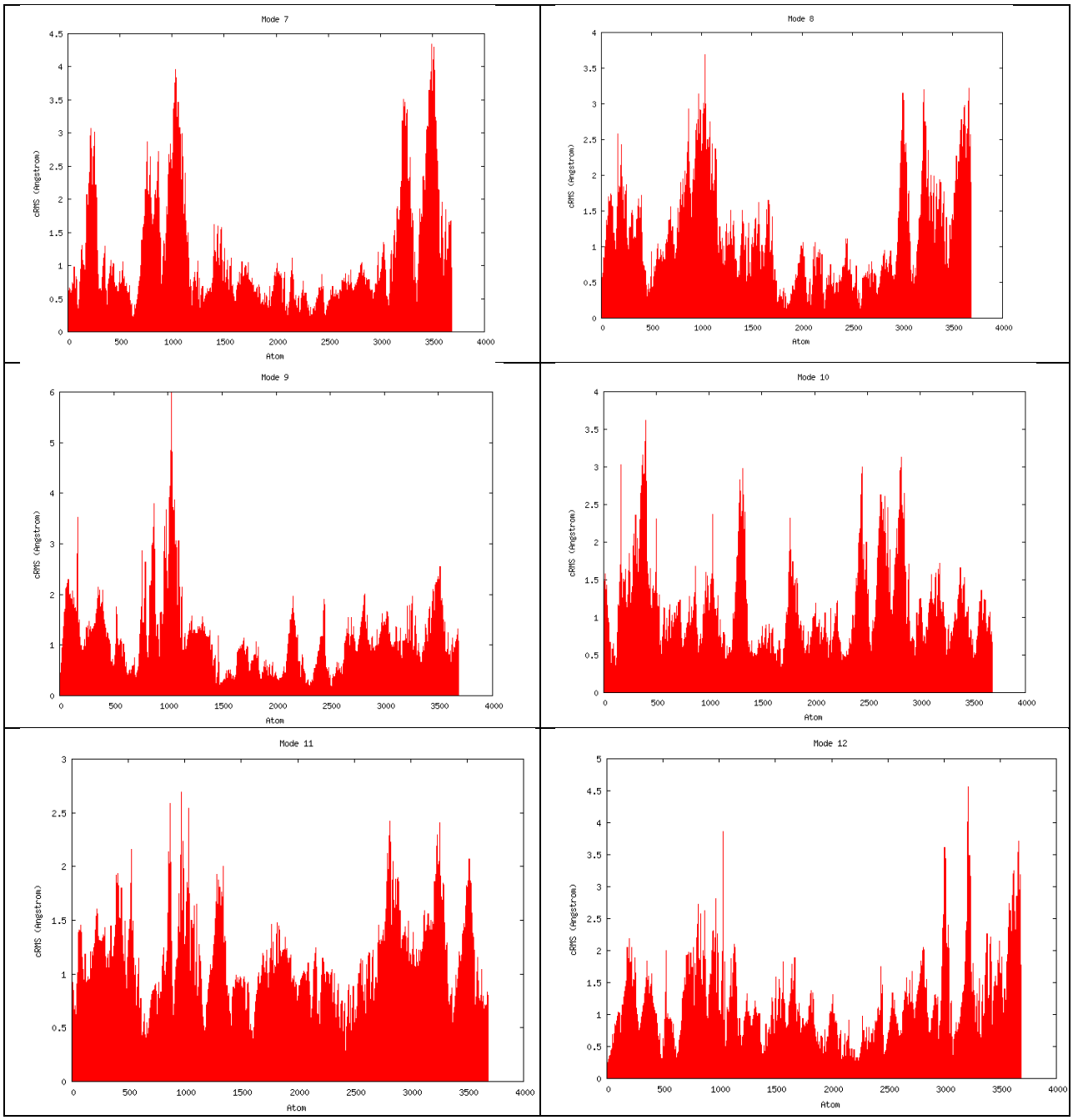


CC plot:



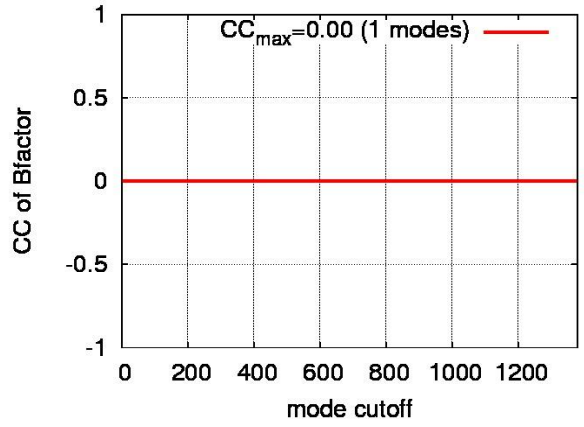
Normal mode analysis:

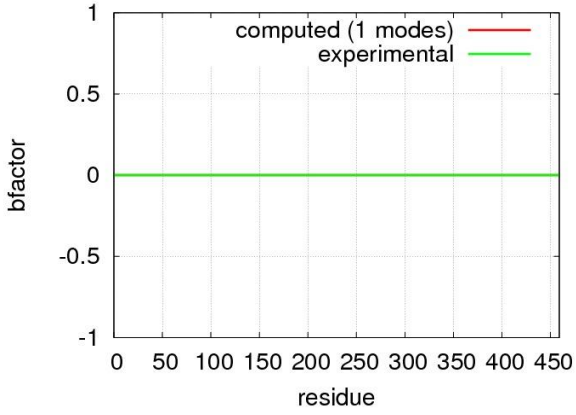



Mode 7
cRMS (Angstrom)
Atom
Mode 8
cRMS (Angstrom)
Atom
Mode 9
cRMS (Angstrom)
Atom
Mode 10
cRMS (Angstrom)
Atom
Mode 11
cRMS (Angstrom)
Atom
Mode 12
cRMS (Angstrom)
Atom

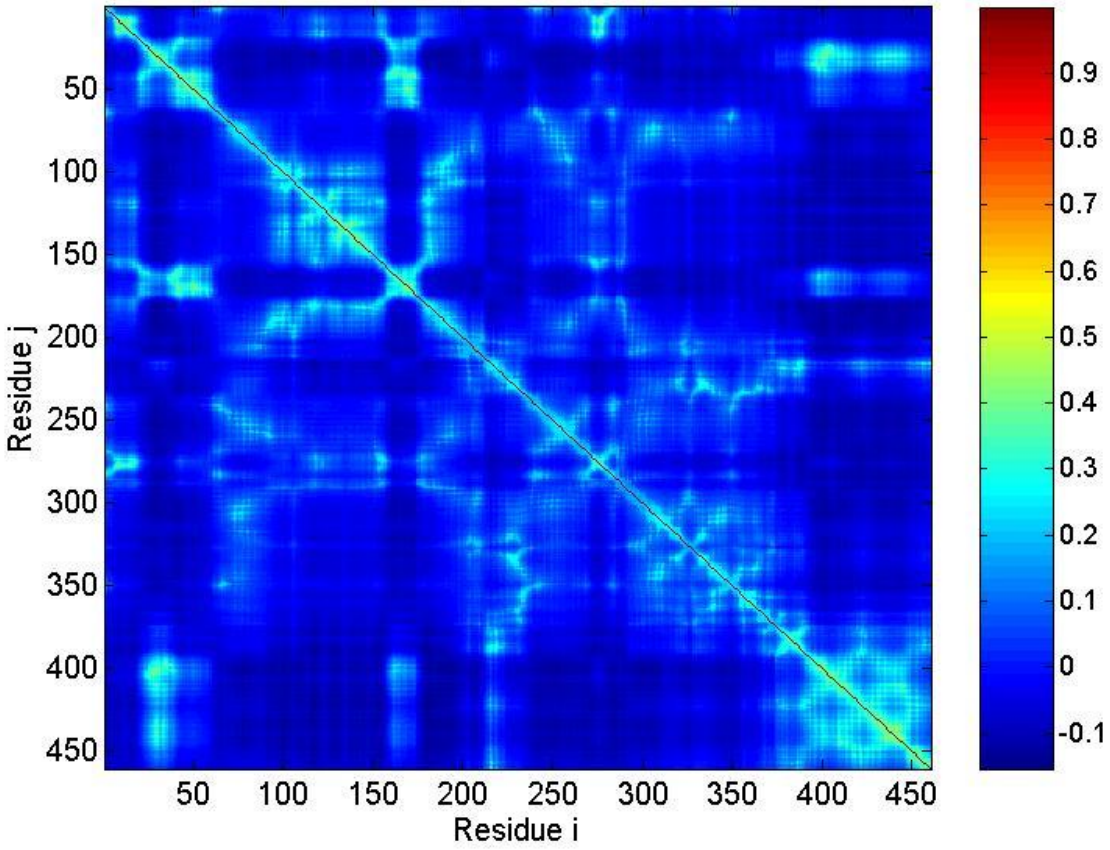
Strictinin- Rhinovirus complex:

Elastic normal mode:

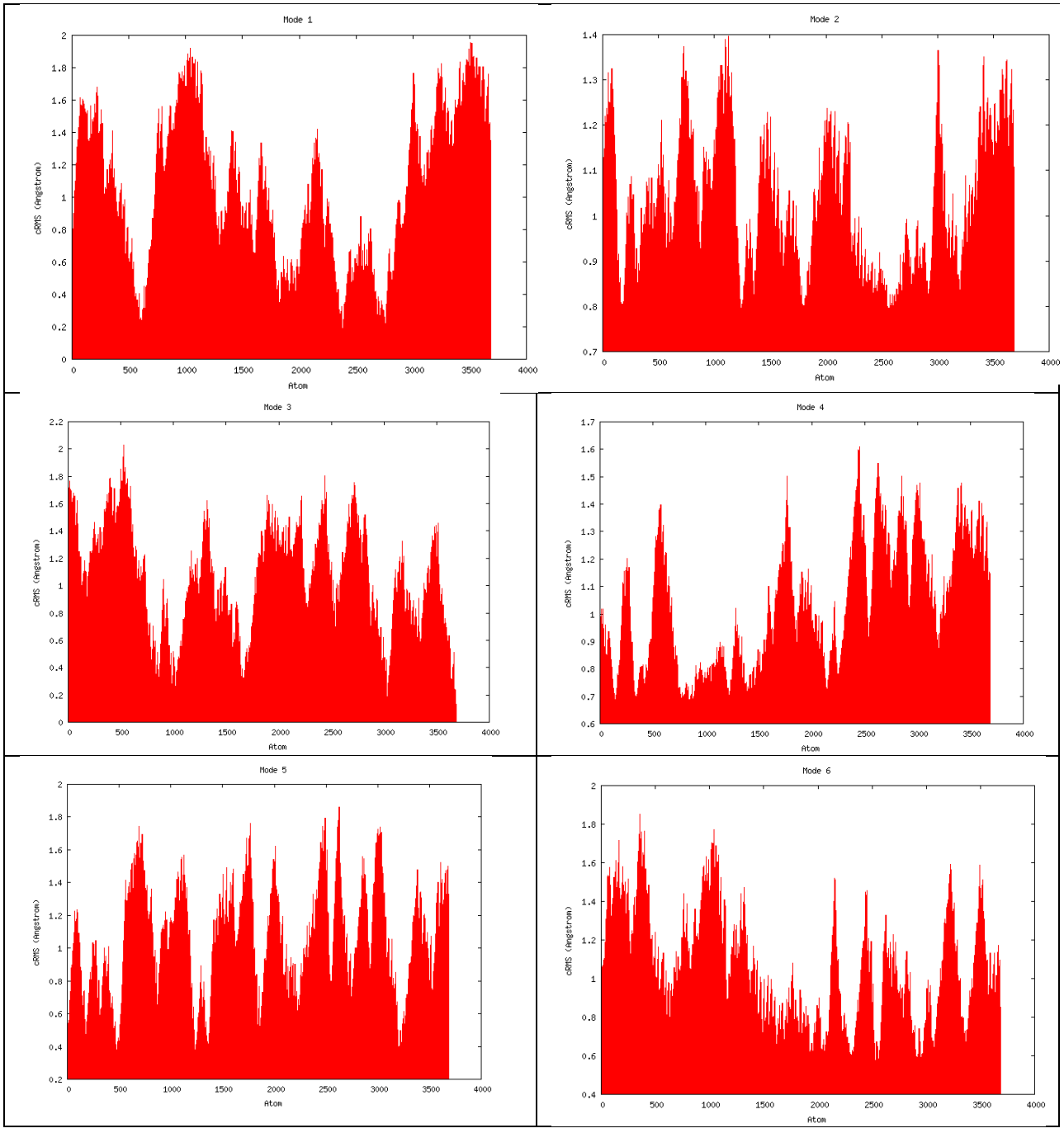


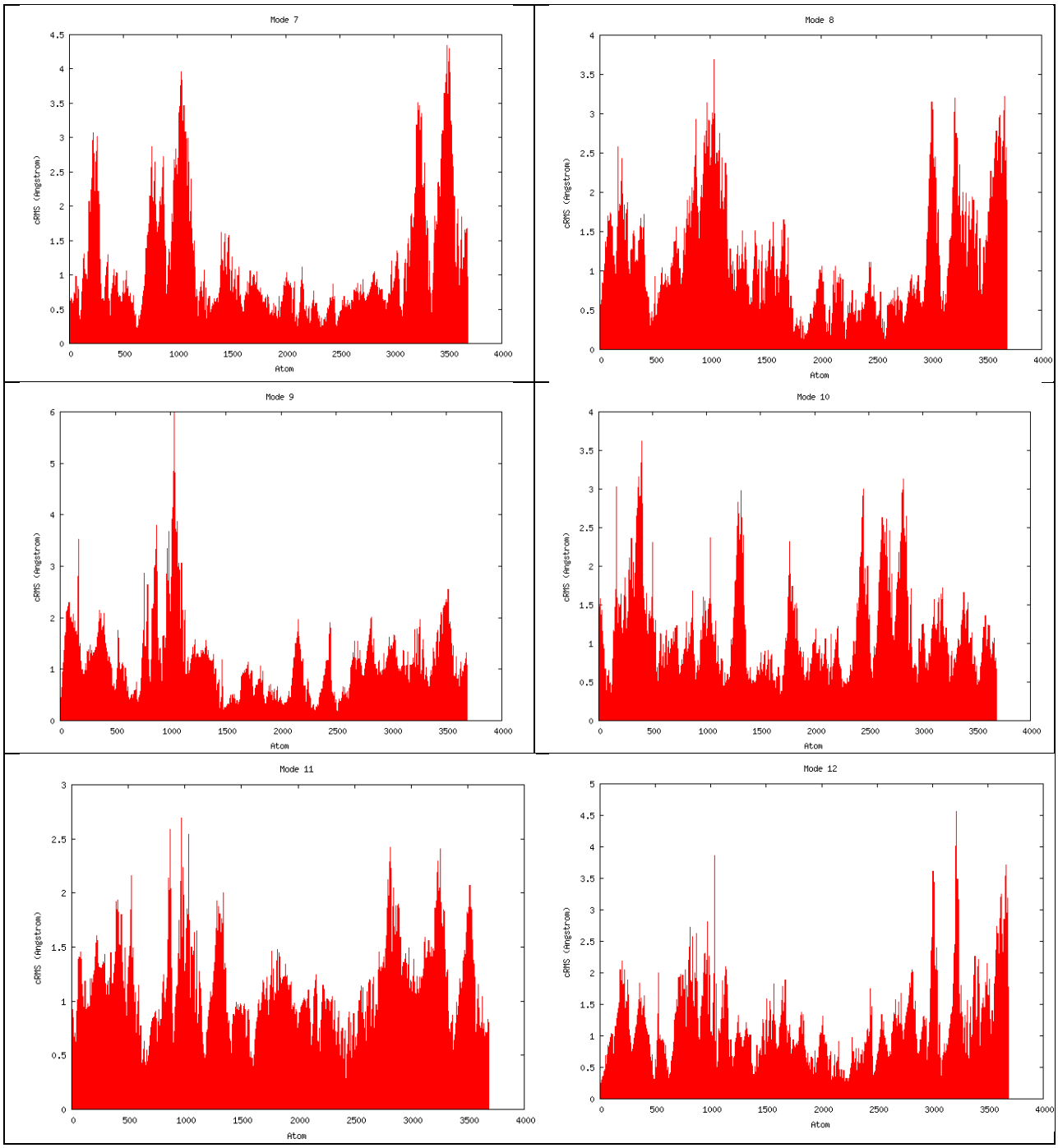


CC plot:



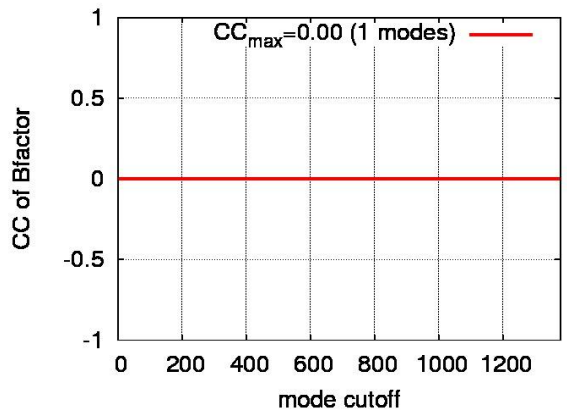
Normal Mode analysis:

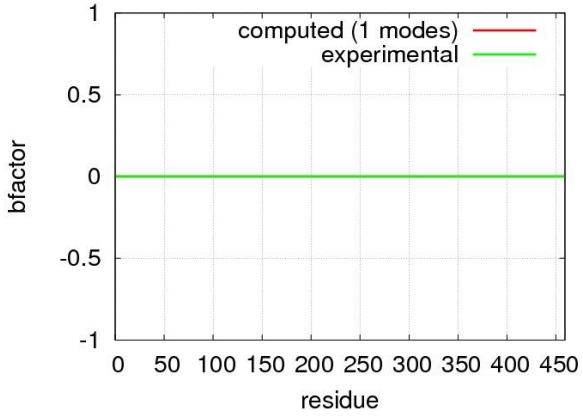



Mode 7
cRMS (Angstrom)
Atom
Mode 8
cRMS (Angstrom)
Atom
Mode 9
cRMS (Angstrom)
Atom
Mode 10
cRMS (Angstrom)
Atom
Mode 11
cRMS (Angstrom)
Atom
Mode 12
cRMS (Angstrom)
Atom

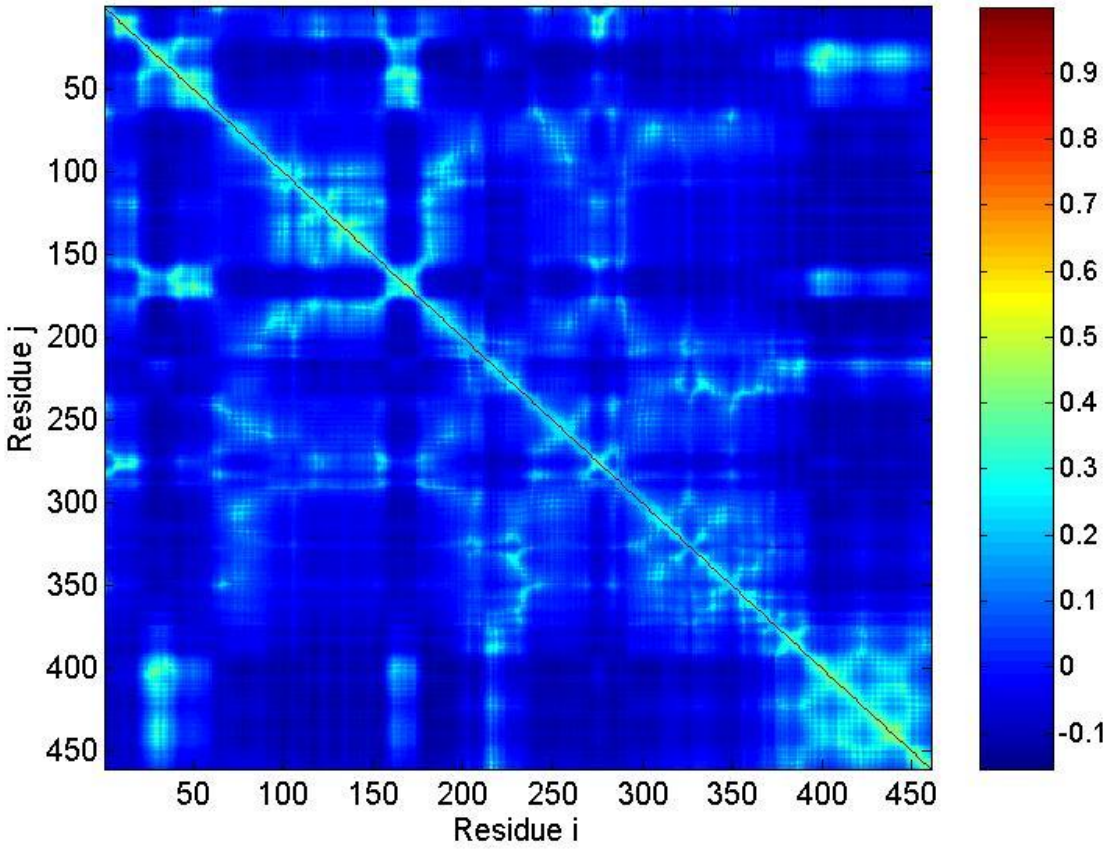
D-glucoside- Rhinovirus complex:

Elastic Normal Mode:

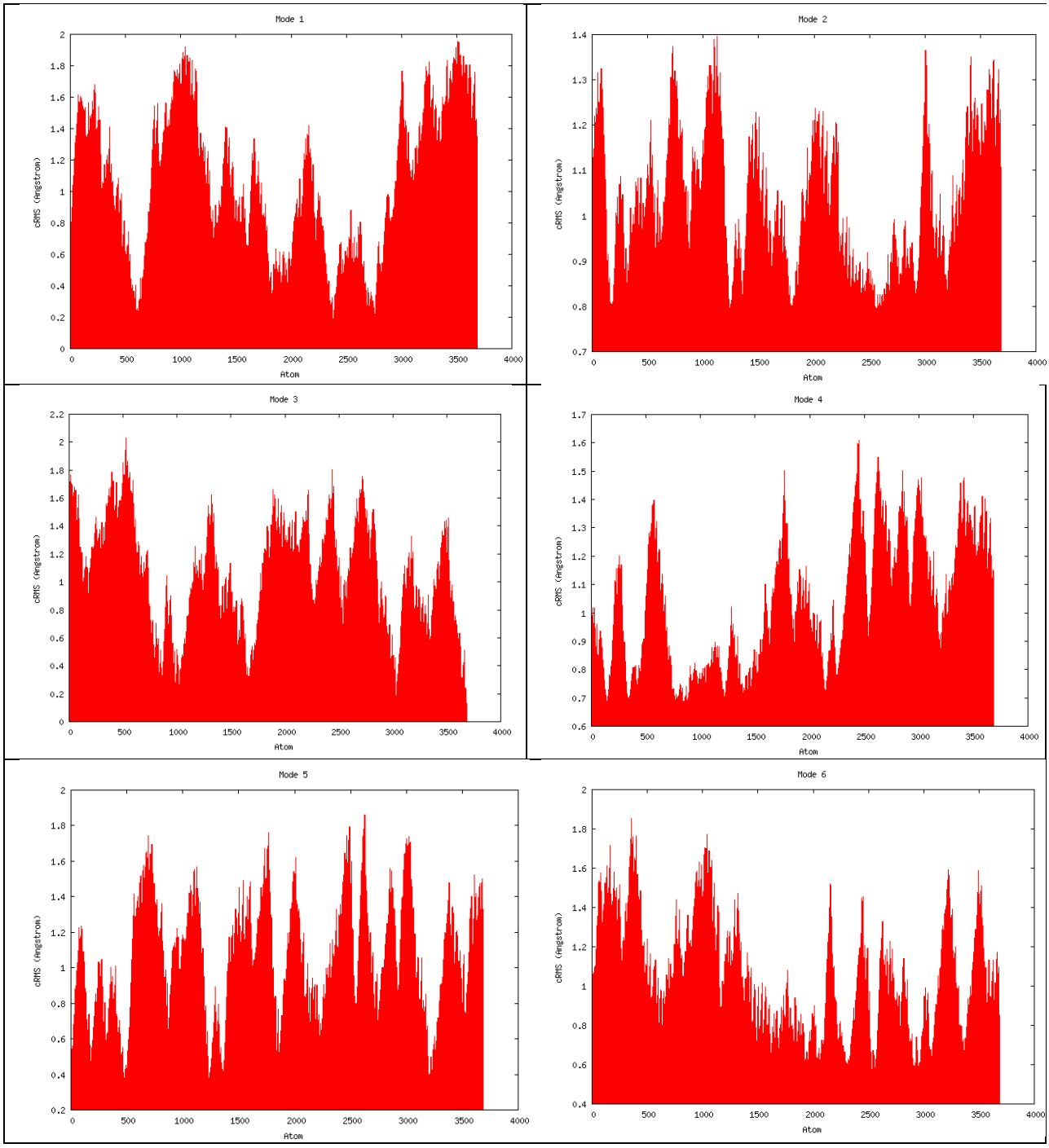


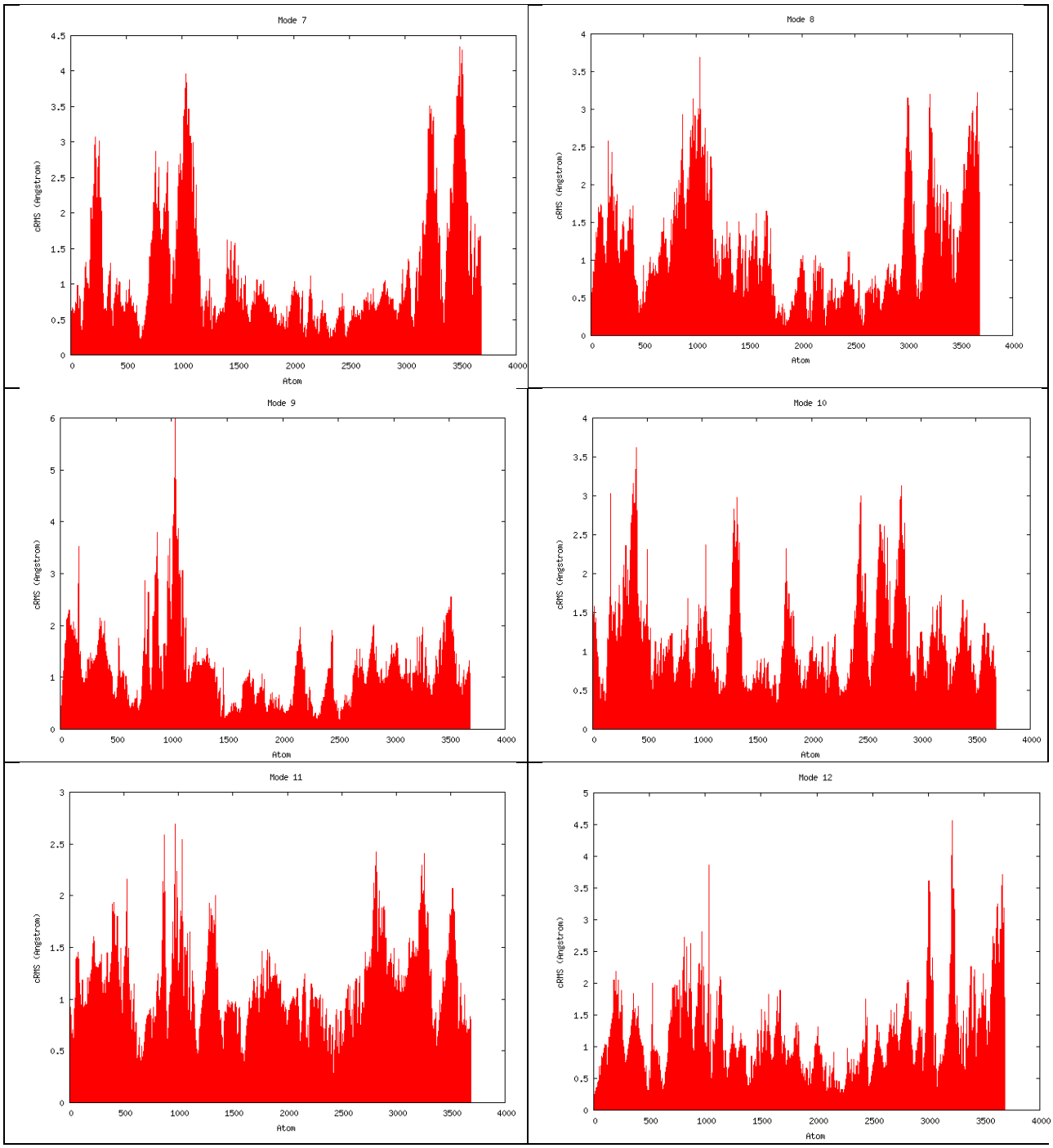


CC plot:



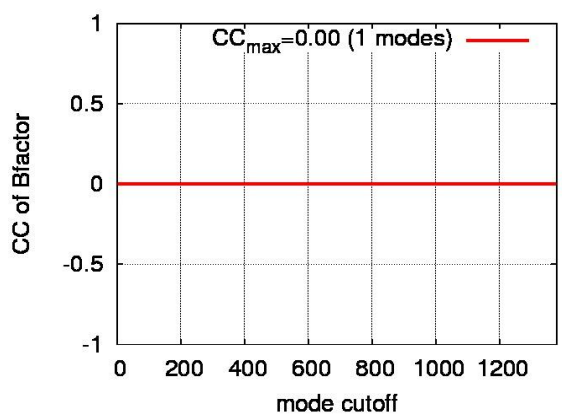
Normal mode analysis:

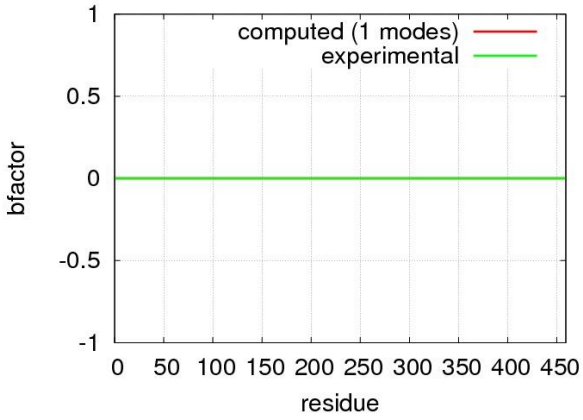



Mode 7
cRMS (Angstrom)
Atom
Mode 8
cRMS (Angstrom)
Atom
Mode 9
cRMS (Angstrom)
Atom
Mode 10
cRMS (Angstrom)
Atom
Mode 11
cRMS (Angstrom)
Atom
Mode 12
cRMS (Angstrom)
Atom

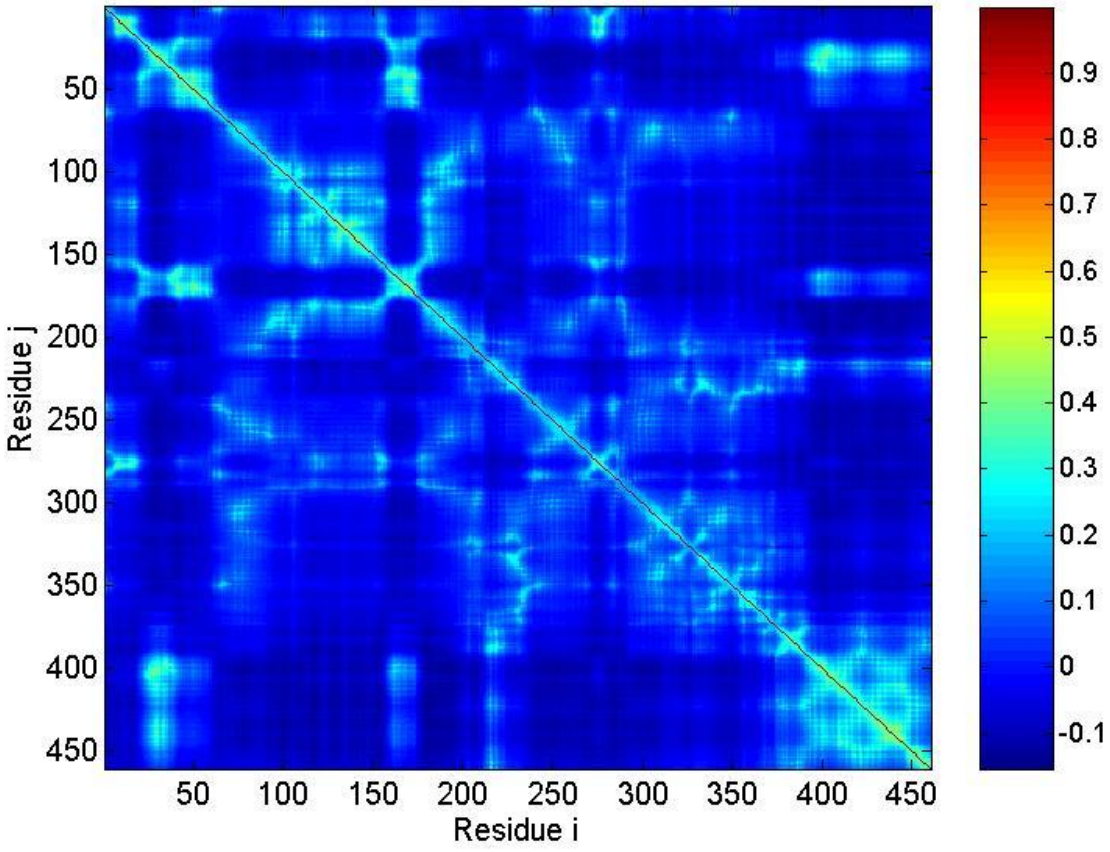
Pinocembrin- Rhinovirus complex:

Elastic normal mode:

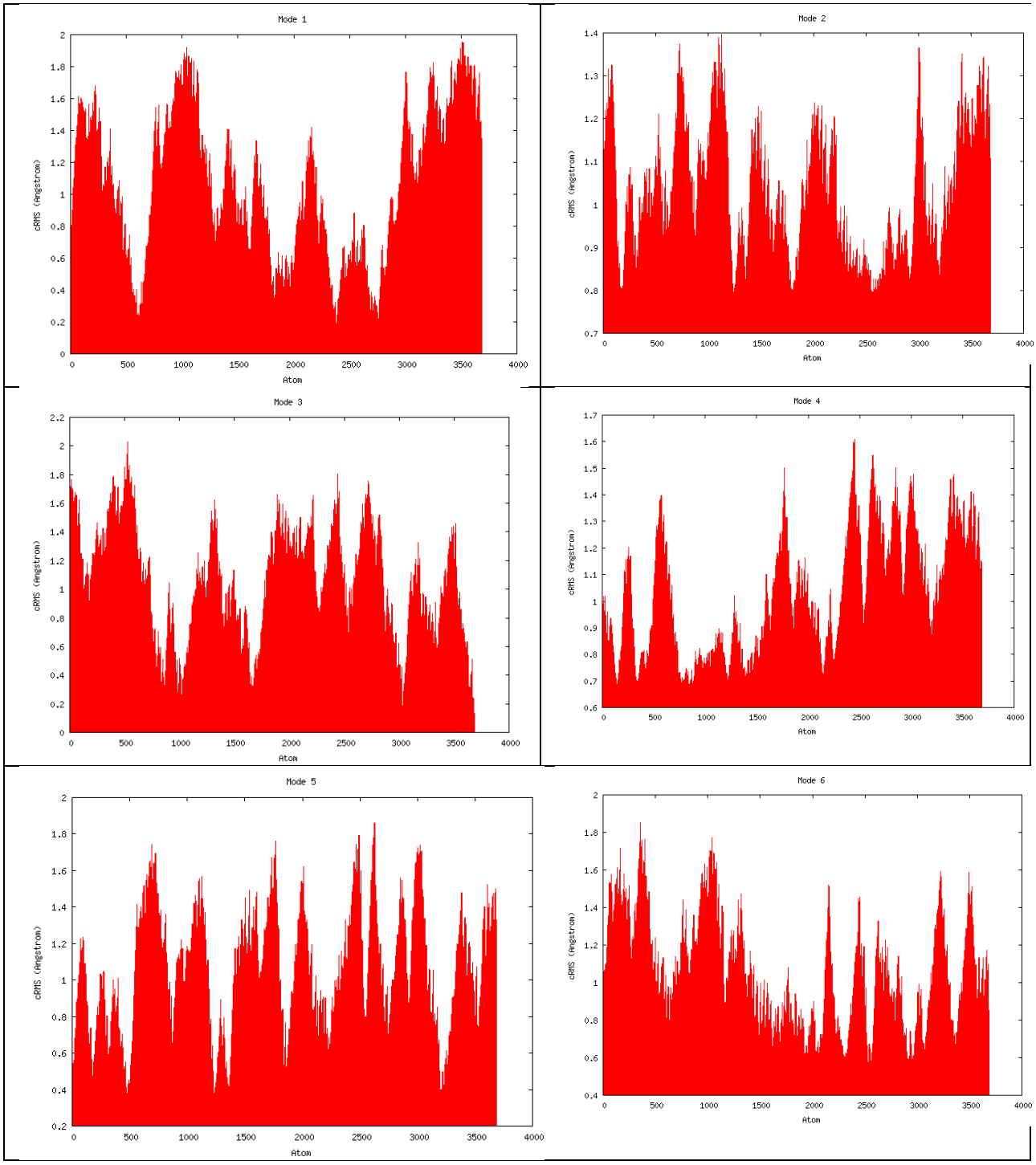


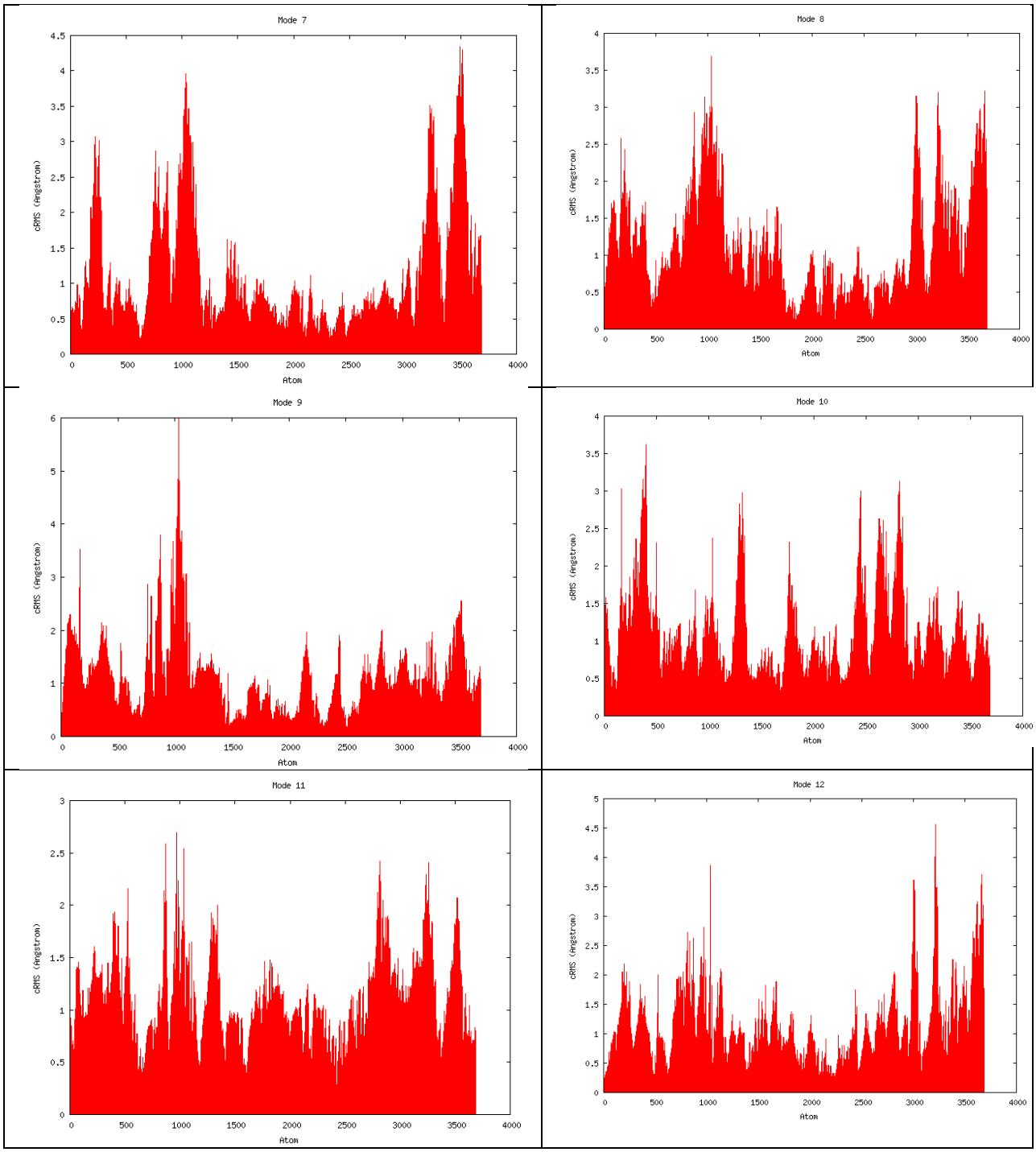


CC plot:



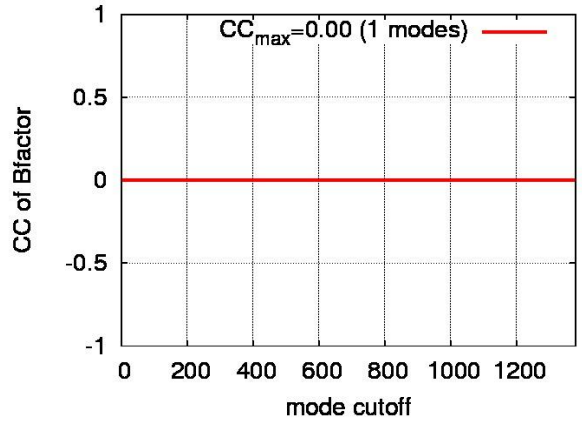
Normal Mode analysis:

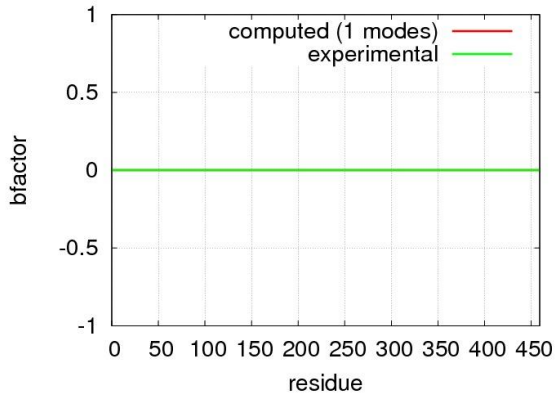



Mode 7
cRMS (Angstrom)
Atom
Mode 8
cRMS (Angstrom)
Atom
Mode 9
cRMS (Angstrom)
Atom
Mode 10
cRMS (Angstrom)
Atom
Mode 11
cRMS (Angstrom)
Atom
Mode 12
cRMS (Angstrom)
Atom

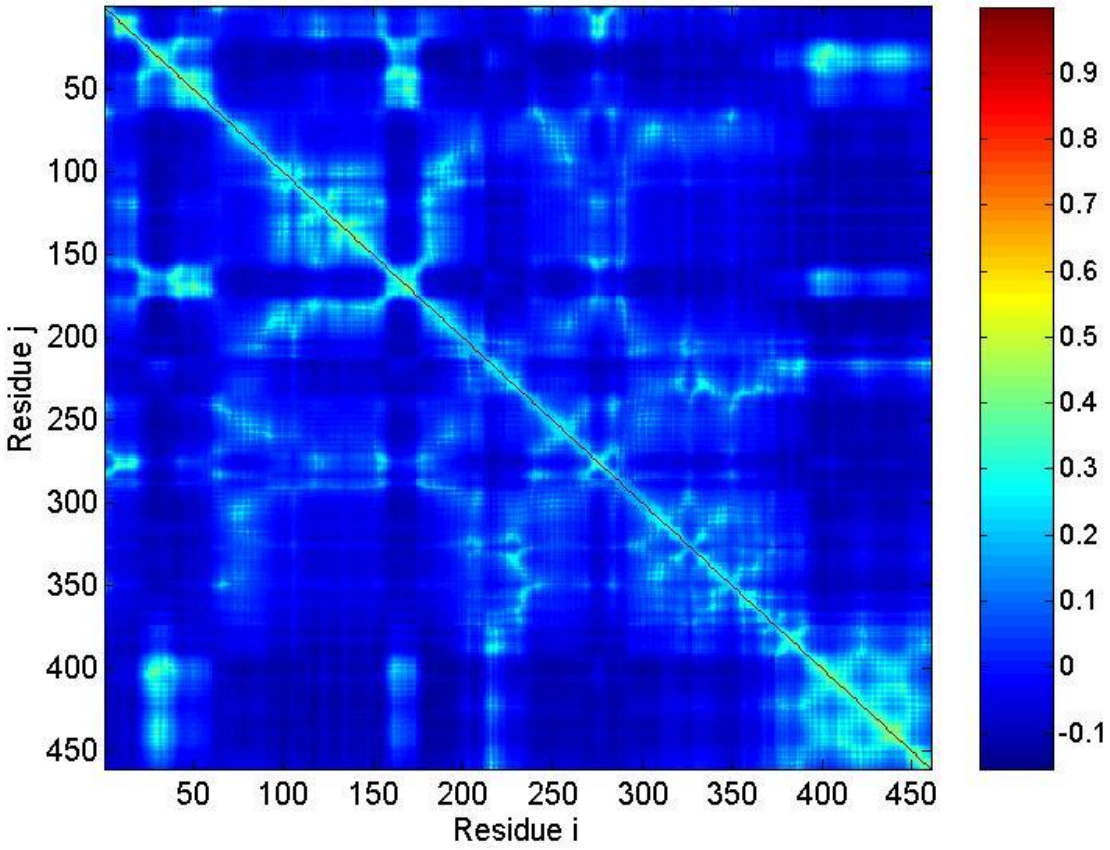
Quercitrin- Rhinovirus complex:

Elastic normal mode:

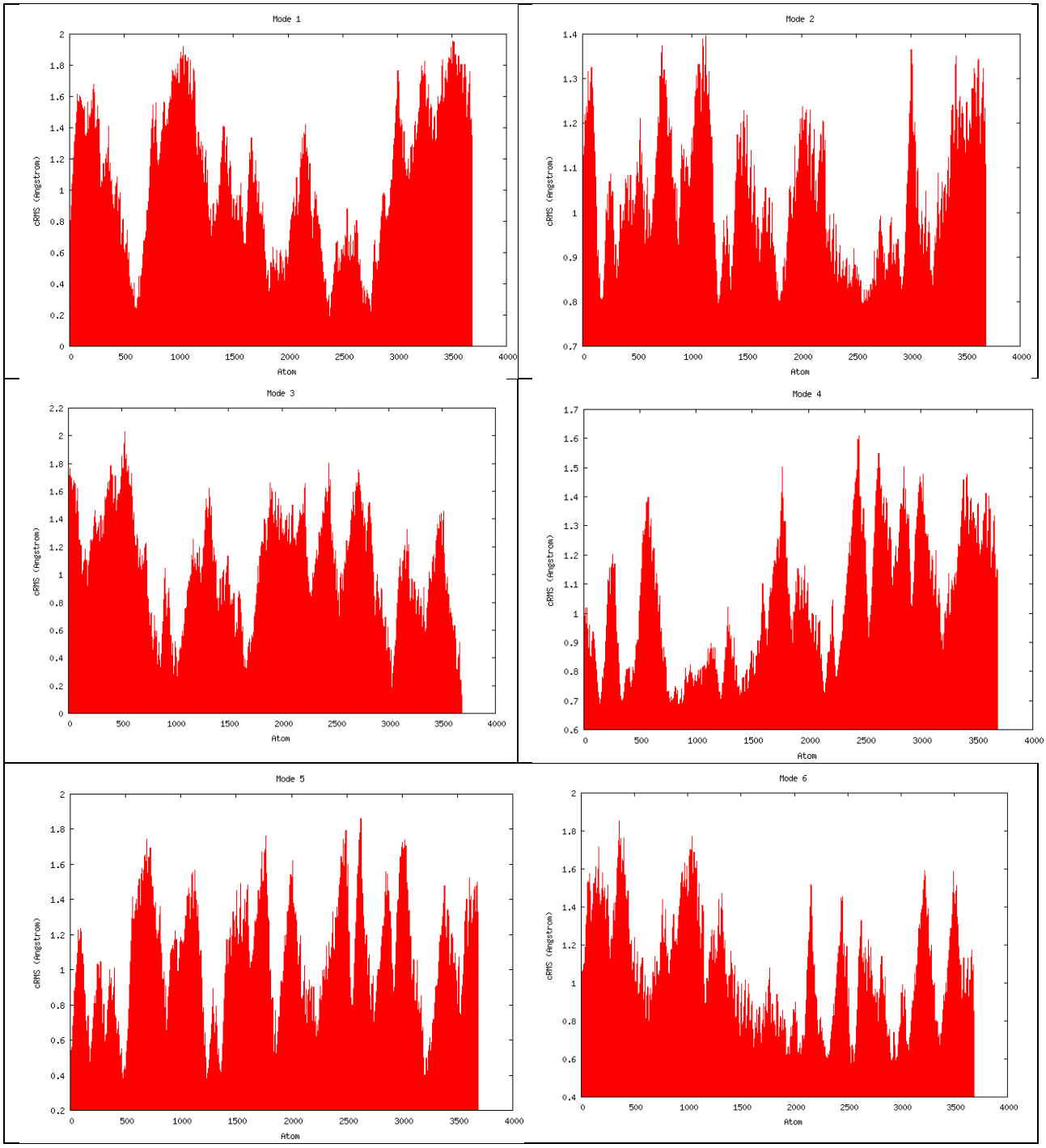


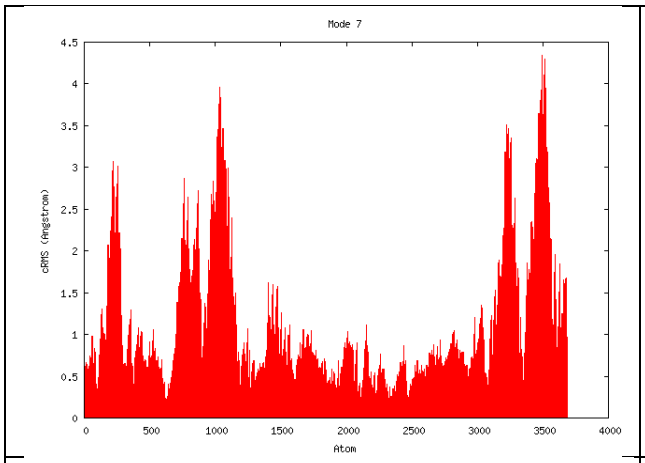


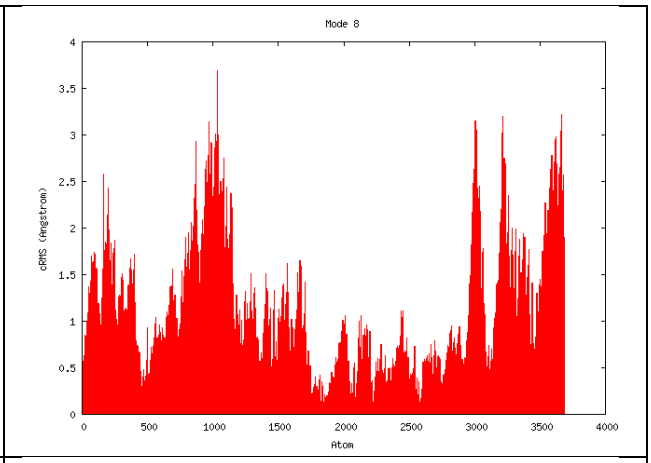
CC Plot:

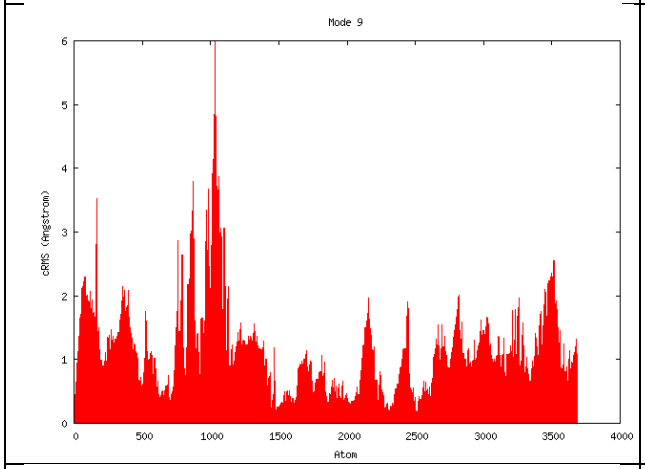


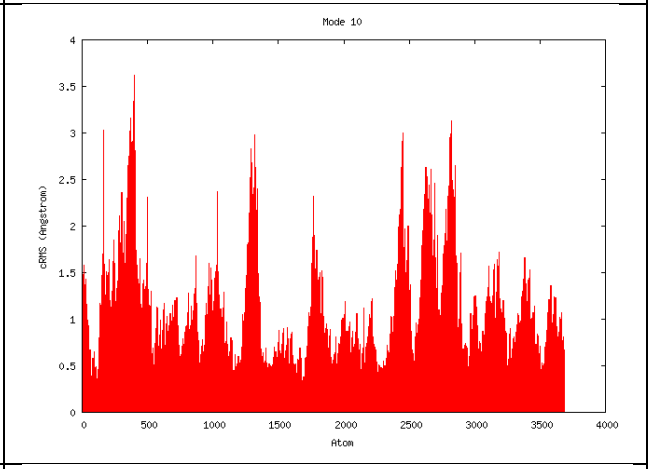
Normal mode analysis:

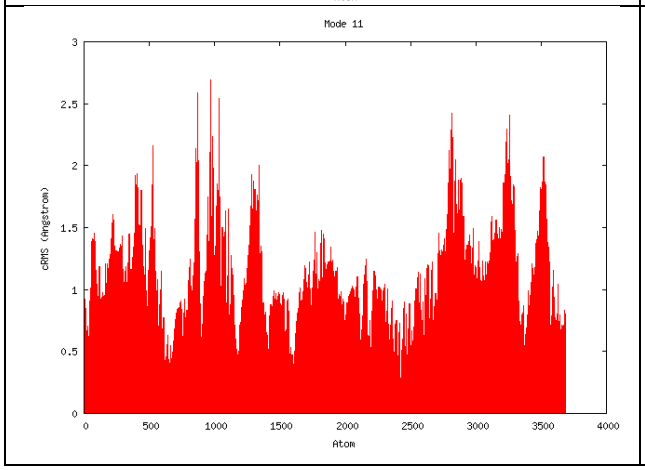


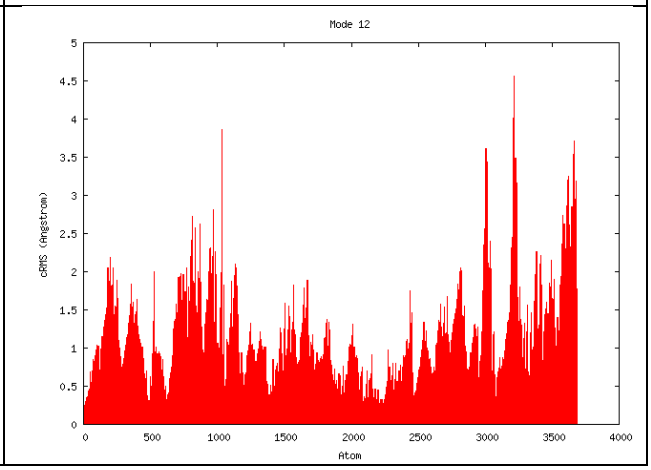

Mode 7
cRMS (Angstrom)
Atom


Mode 8
cRMS (Angstrom)
Atom


Mode 9
cRMS (Angstrom)
Atom

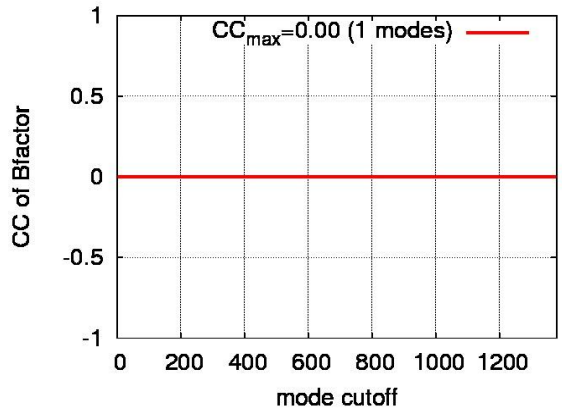

Mode 10
cRMS (Angstrom)
Atom

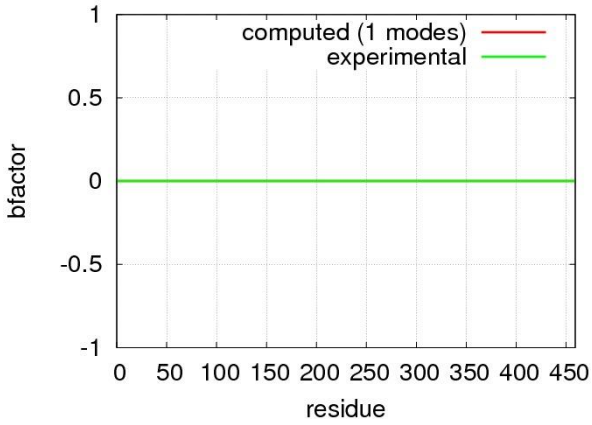

Mode 11
cRMS (Angstrom)
Atom


Mode 12
cRMS (Angstrom)
Atom

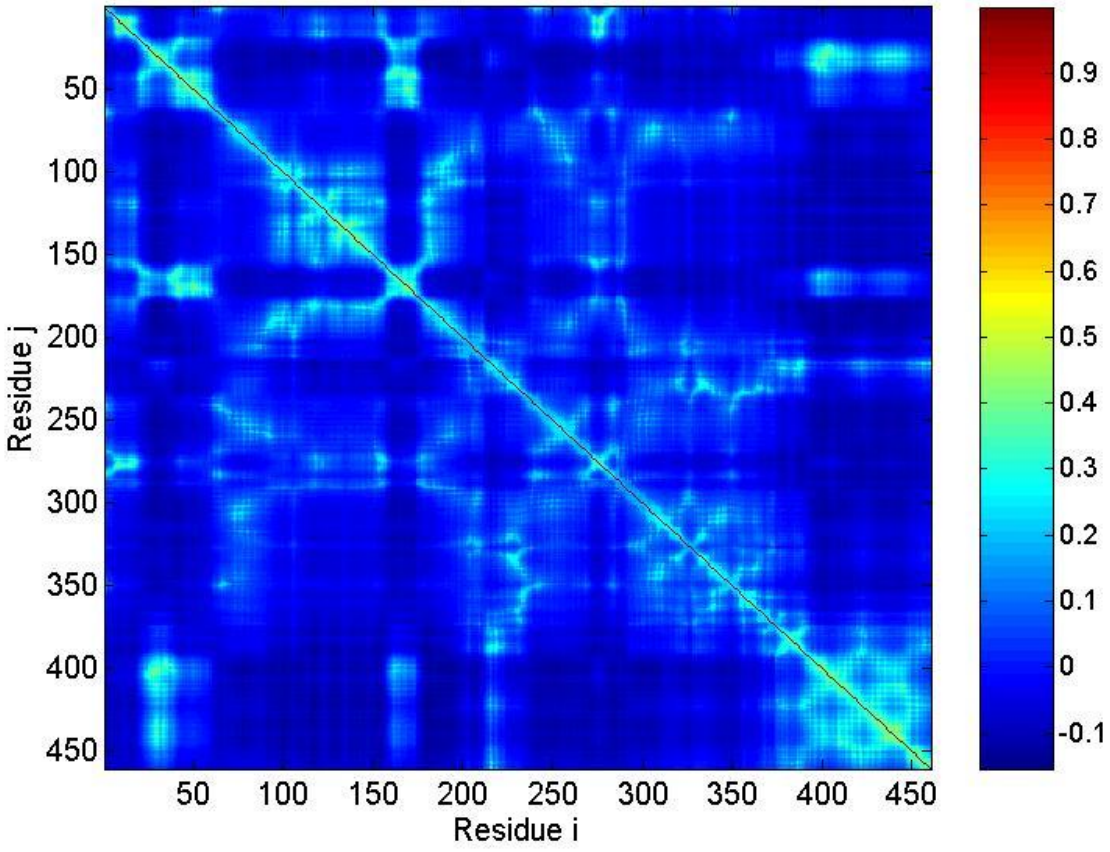
Theaflavin- Rhinovirus complex:

Elastic normal mode:

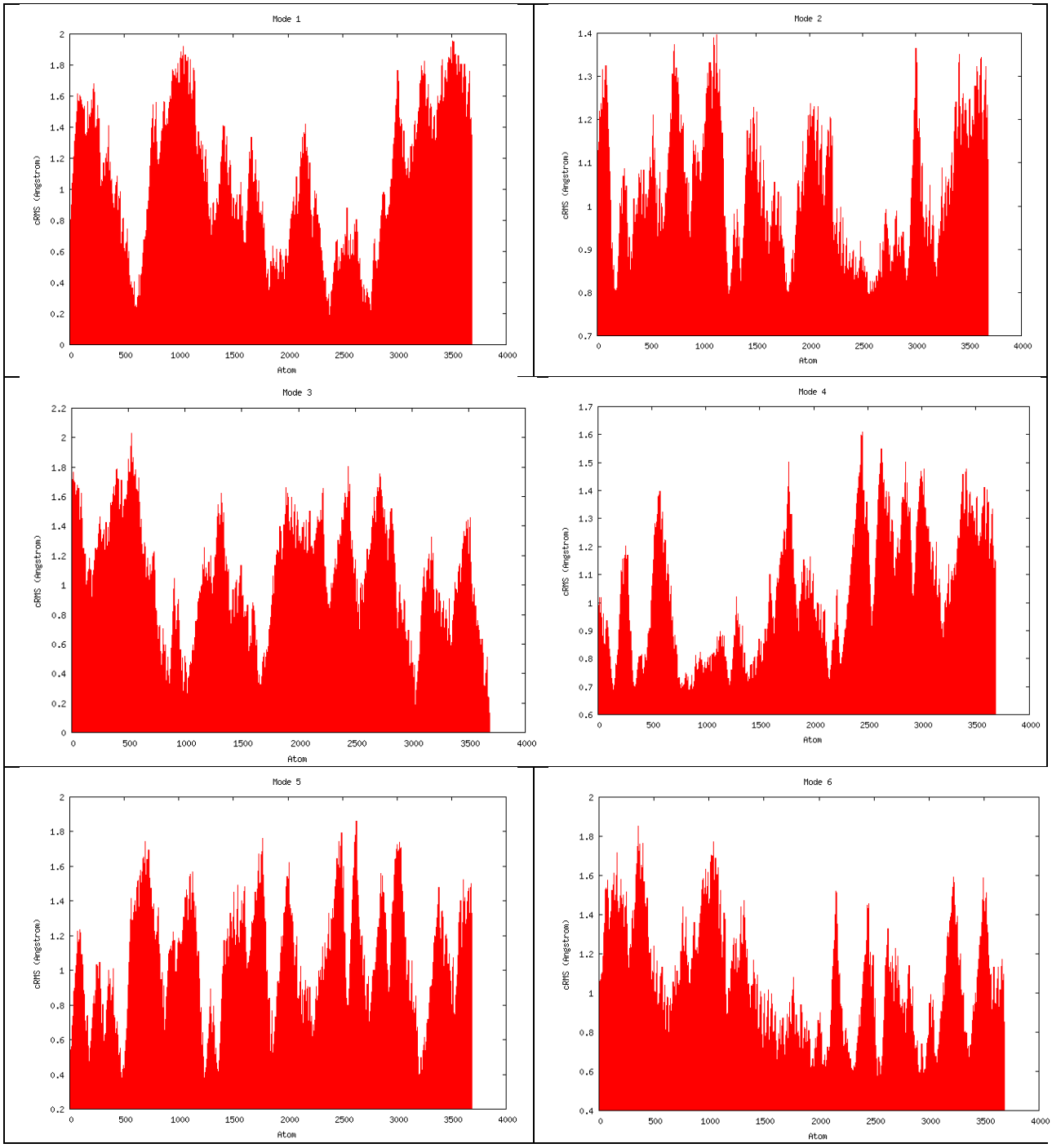


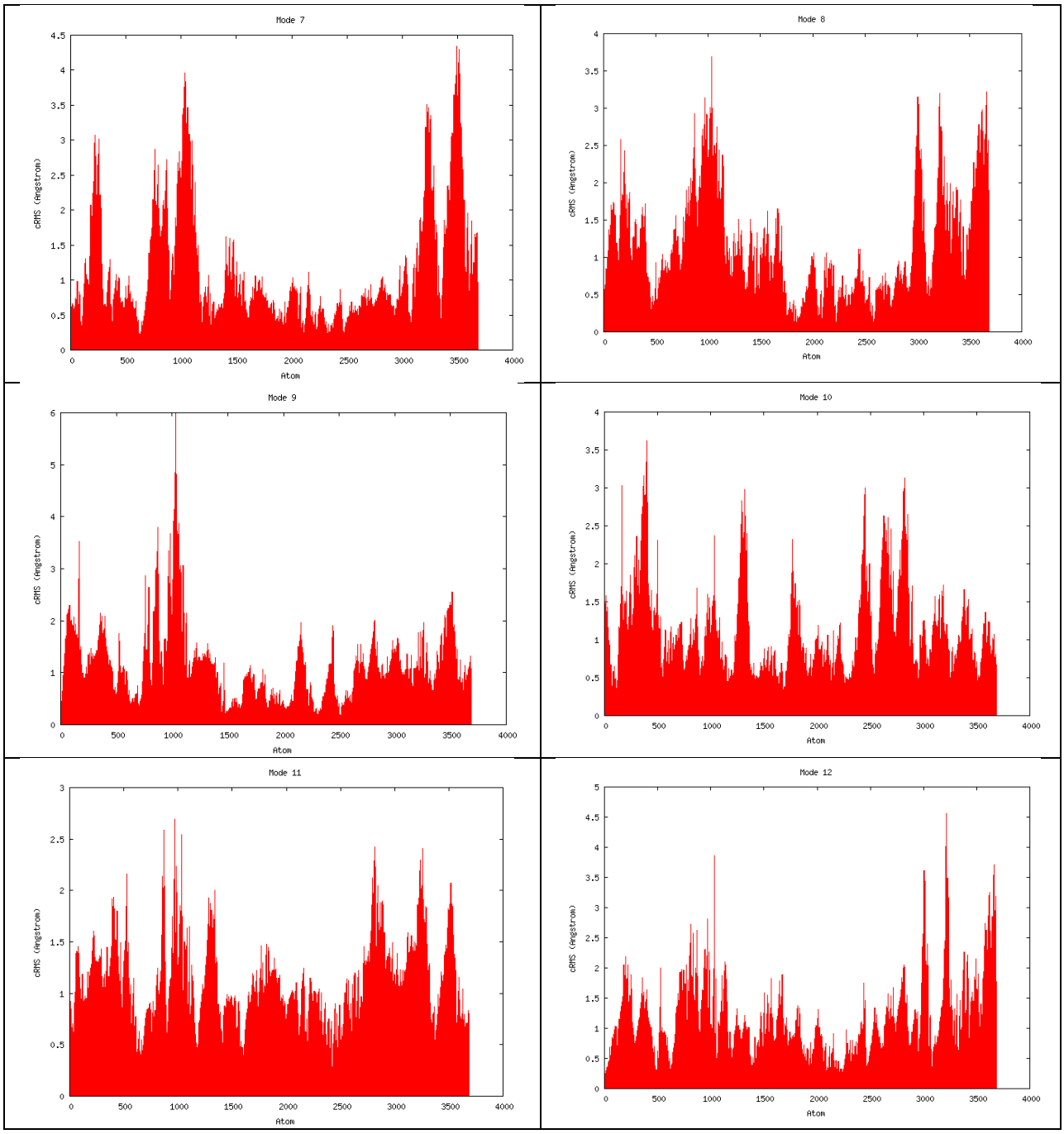


CC Plot:



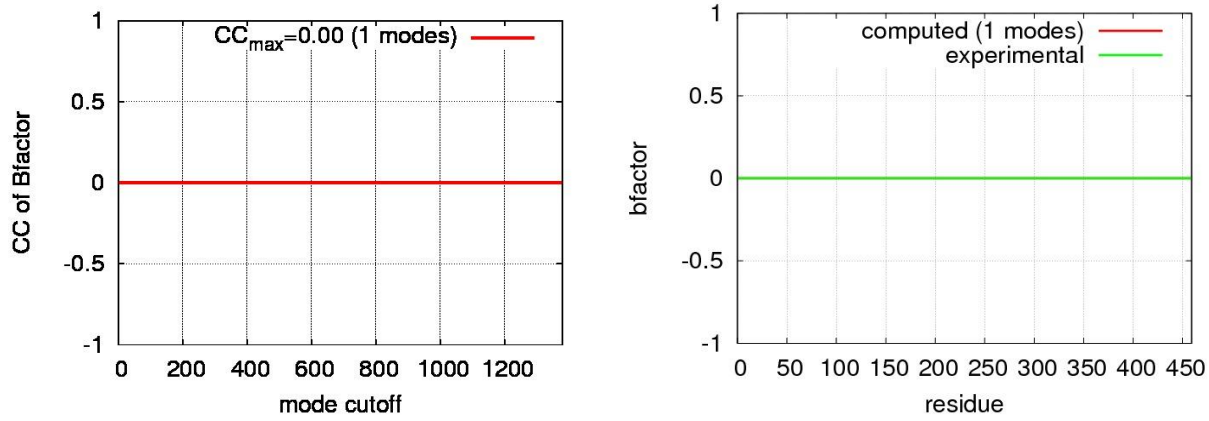
Normal mode analysis:



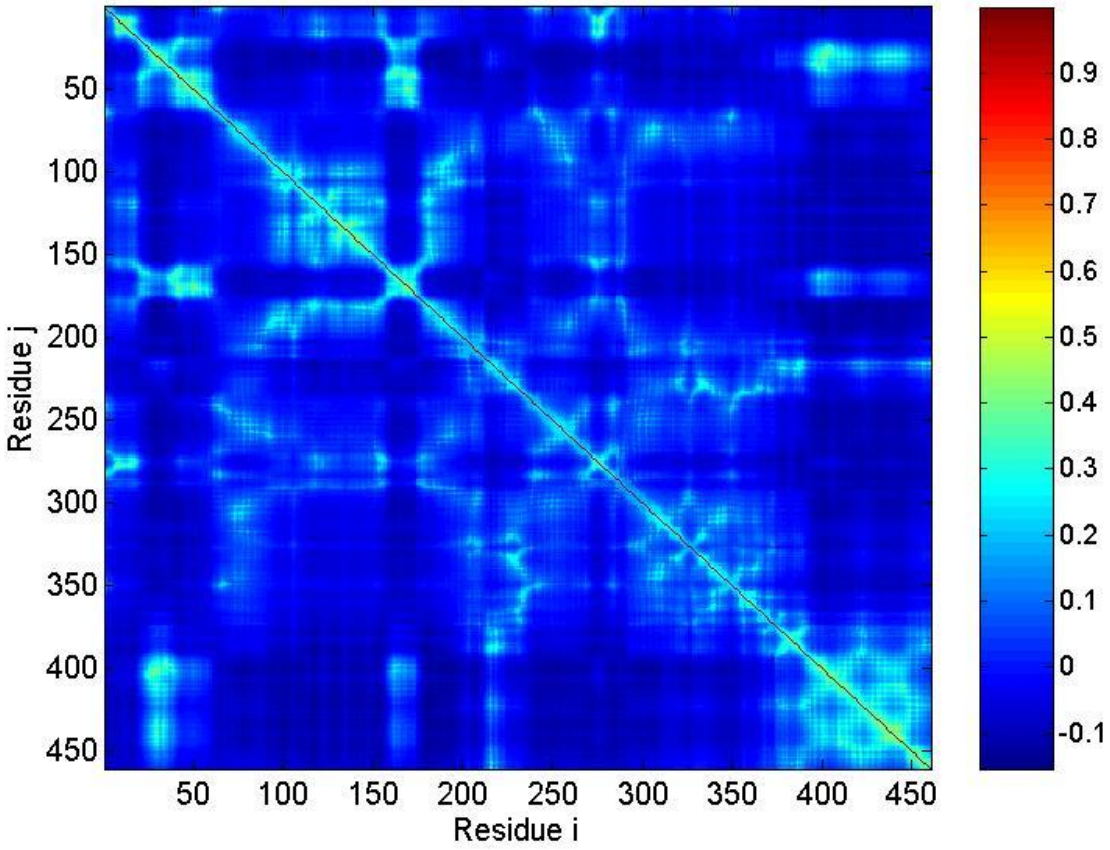

Mode 7
cRMS (Angstrom)
Atom
Mode 8
cRMS (Angstrom)
Atom
Mode 9
cRMS (Angstrom)
Atom
Mode 10
cRMS (Angstrom)
Atom
Mode 11
cRMS (Angstrom)
Atom
Mode 12
cRMS (Angstrom)
Atom

Myricitrin- Rhinovirus complex:

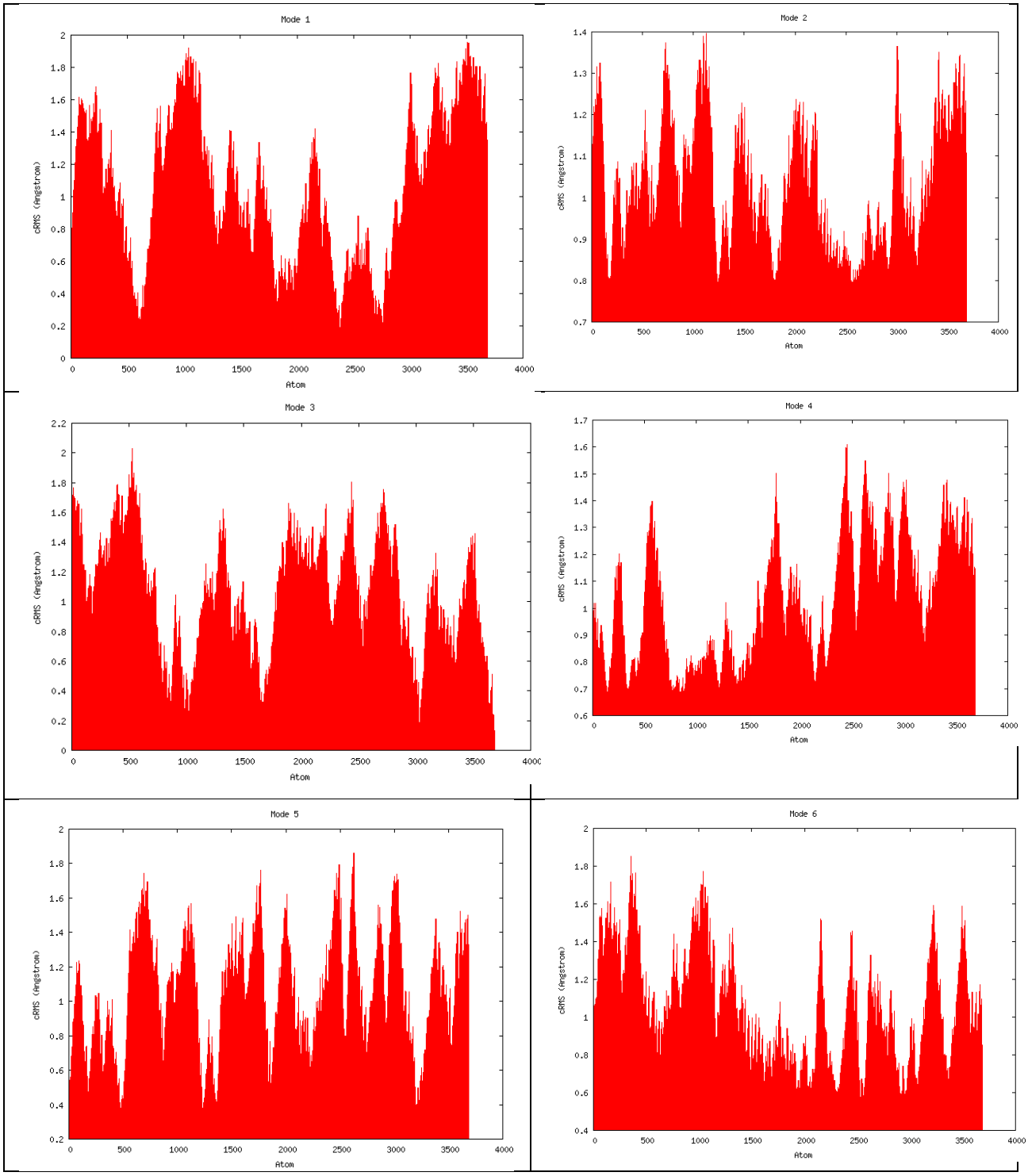
Elastic normal mode:

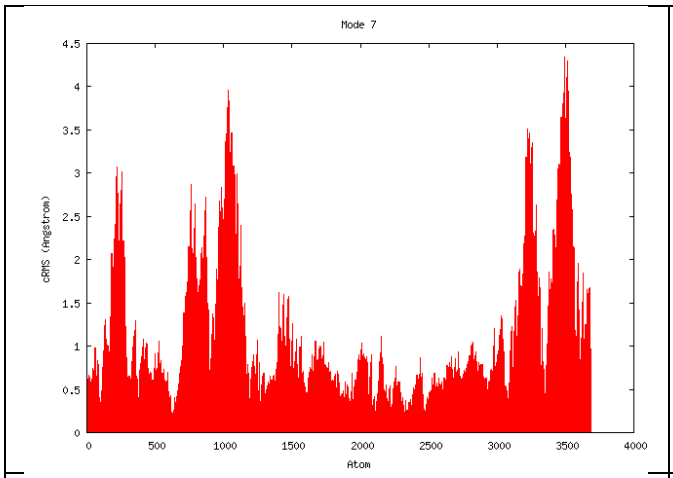


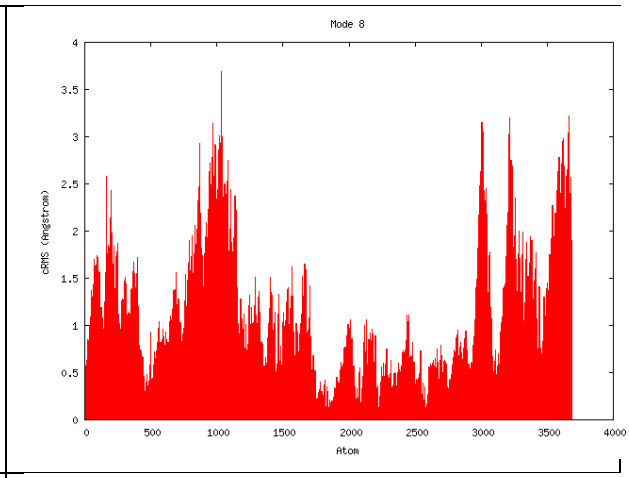
CC Plot:

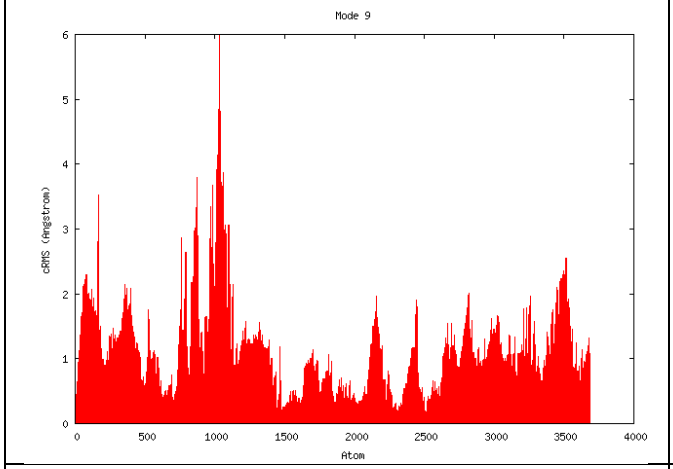


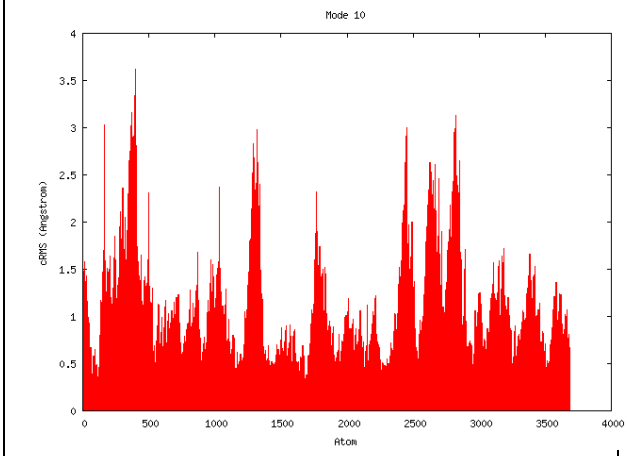
Normal mode analysis:

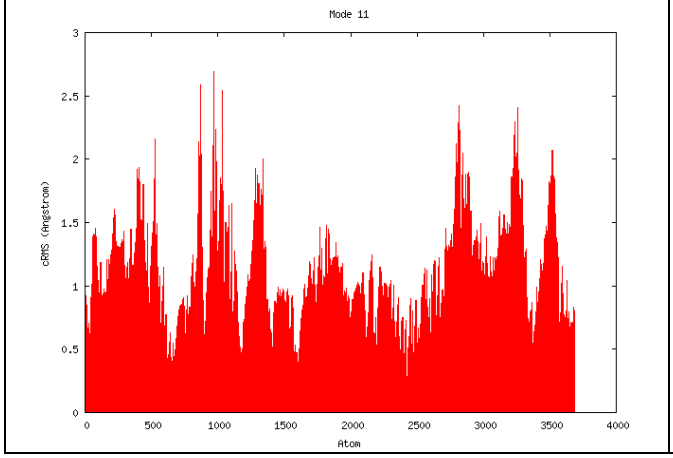


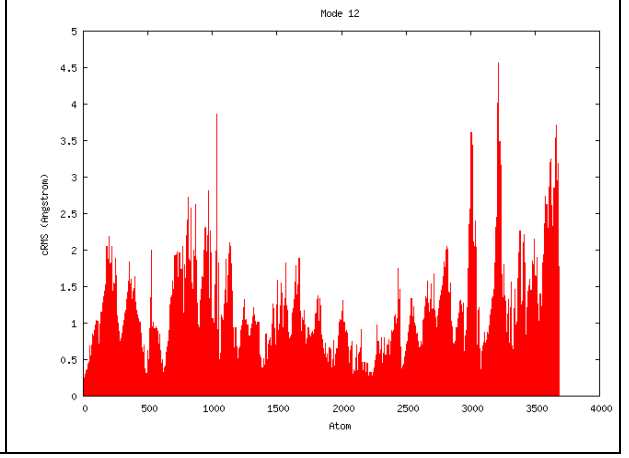

Mode 7
cRMS (Angstrom)
Atom


Mode 8
cRMS (Angstrom)
Atom


Mode 9
cRMS (Angstrom)
Atom

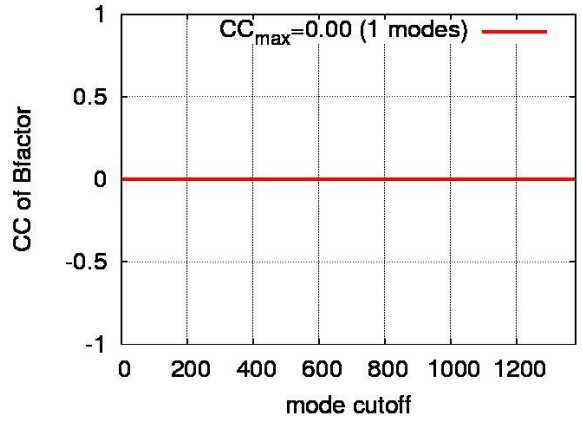

Mode 10
cRMS (Angstrom)
Atom

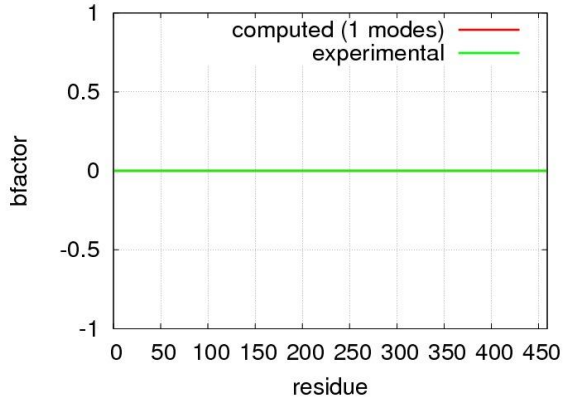

Mode 11
cRMS (Angstrom)
Atom


Mode 12
cRMS (Angstrom)
Atom

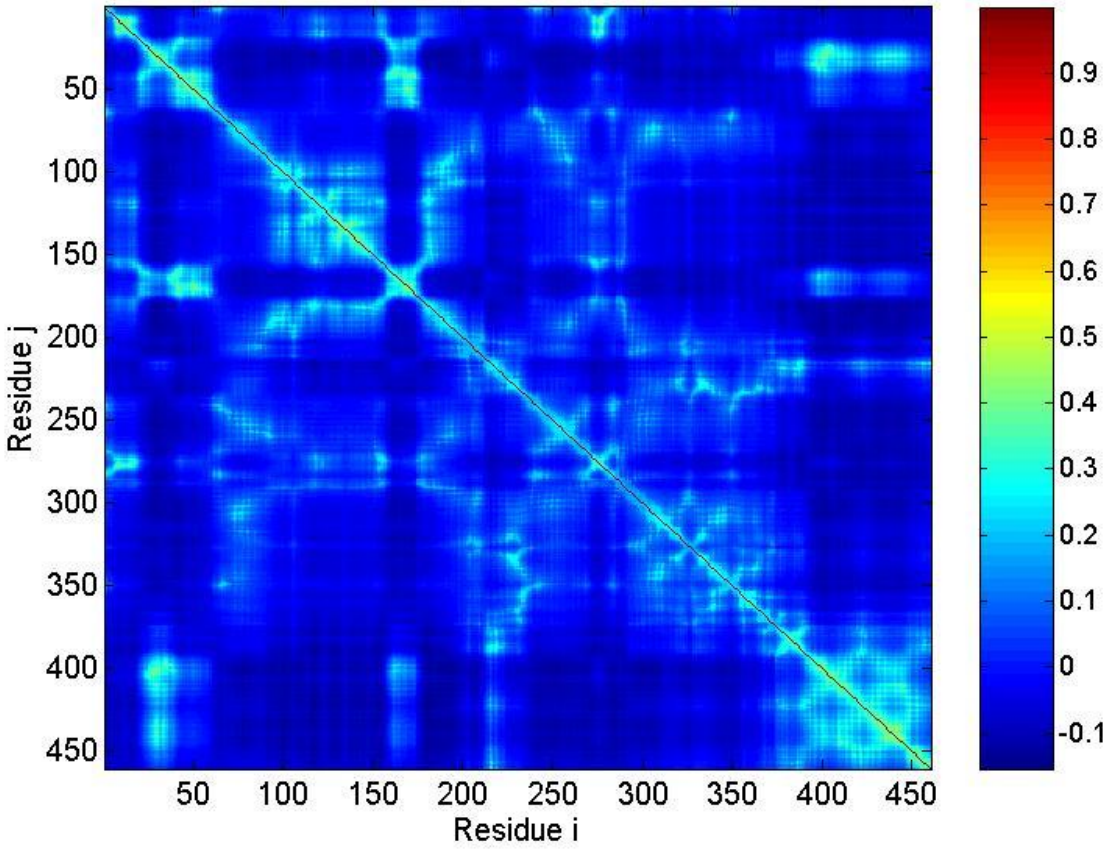
Valine- Rhinovirus complex:

Elastic Normal mode:

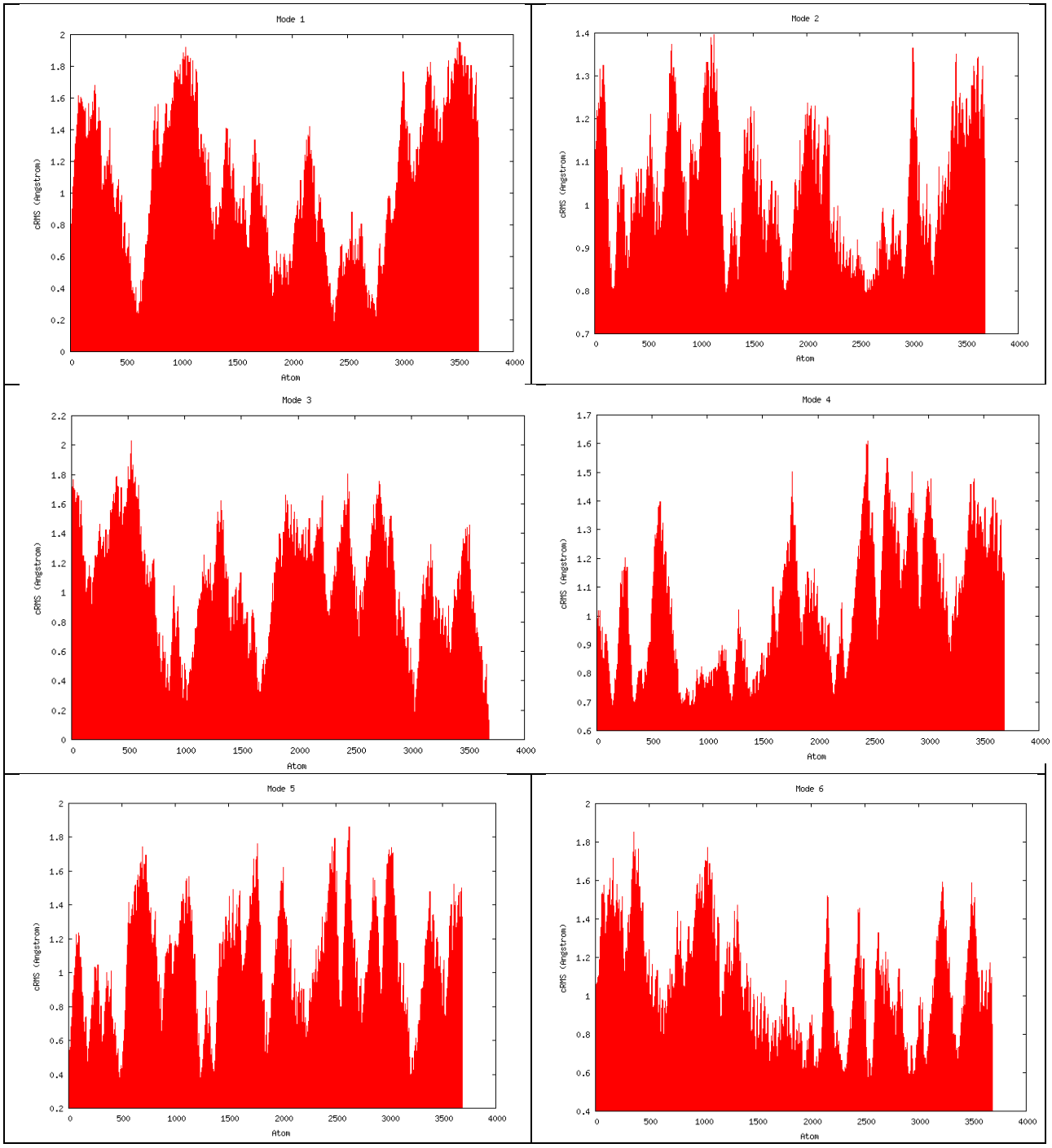


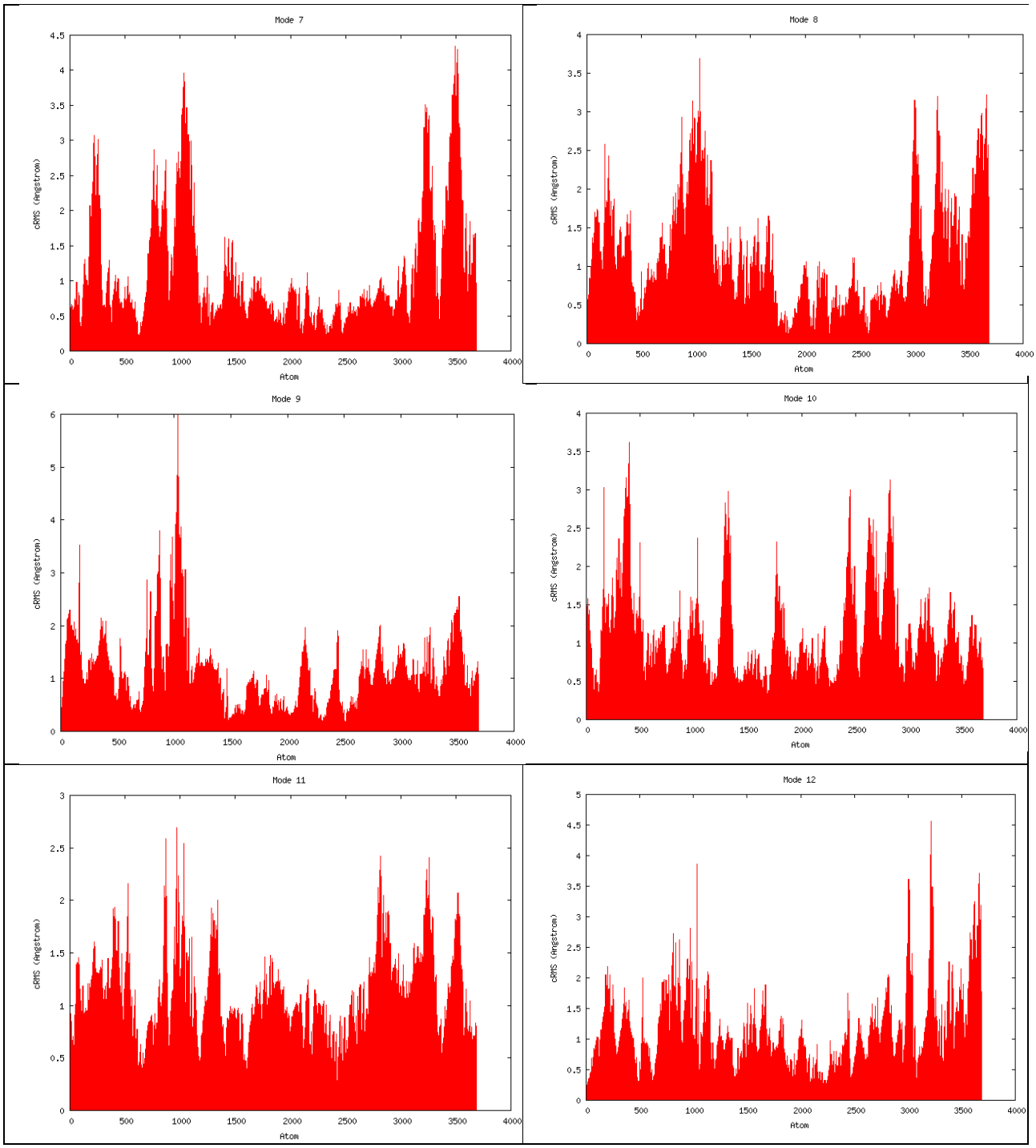


CC Plot:



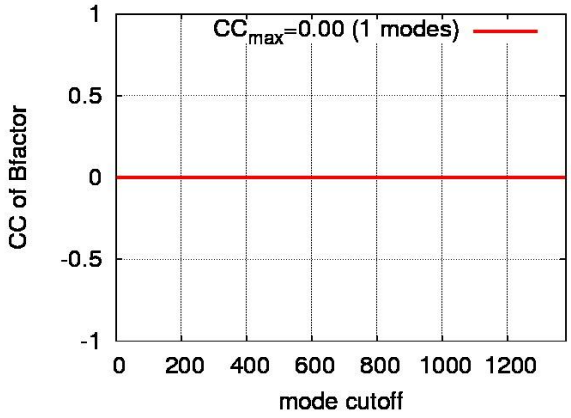
Normal mode analysis:

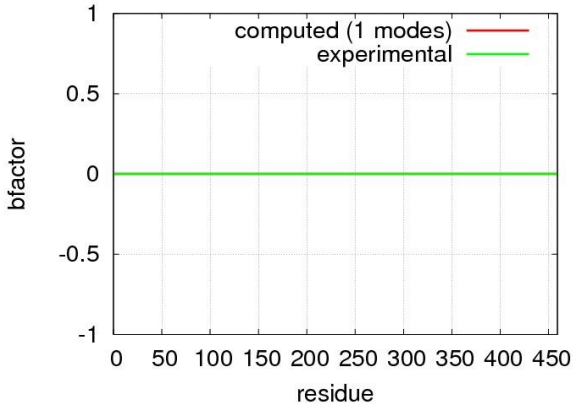



Mode 7
cRMS (Angstrom)
Atom
Mode 8
cRMS (Angstrom)
Atom
Mode 9
cRMS (Angstrom)
Atom
Mode 10
cRMS (Angstrom)
Atom
Mode 11
cRMS (Angstrom)
Atom
Mode 12
cRMS (Angstrom)
Atom

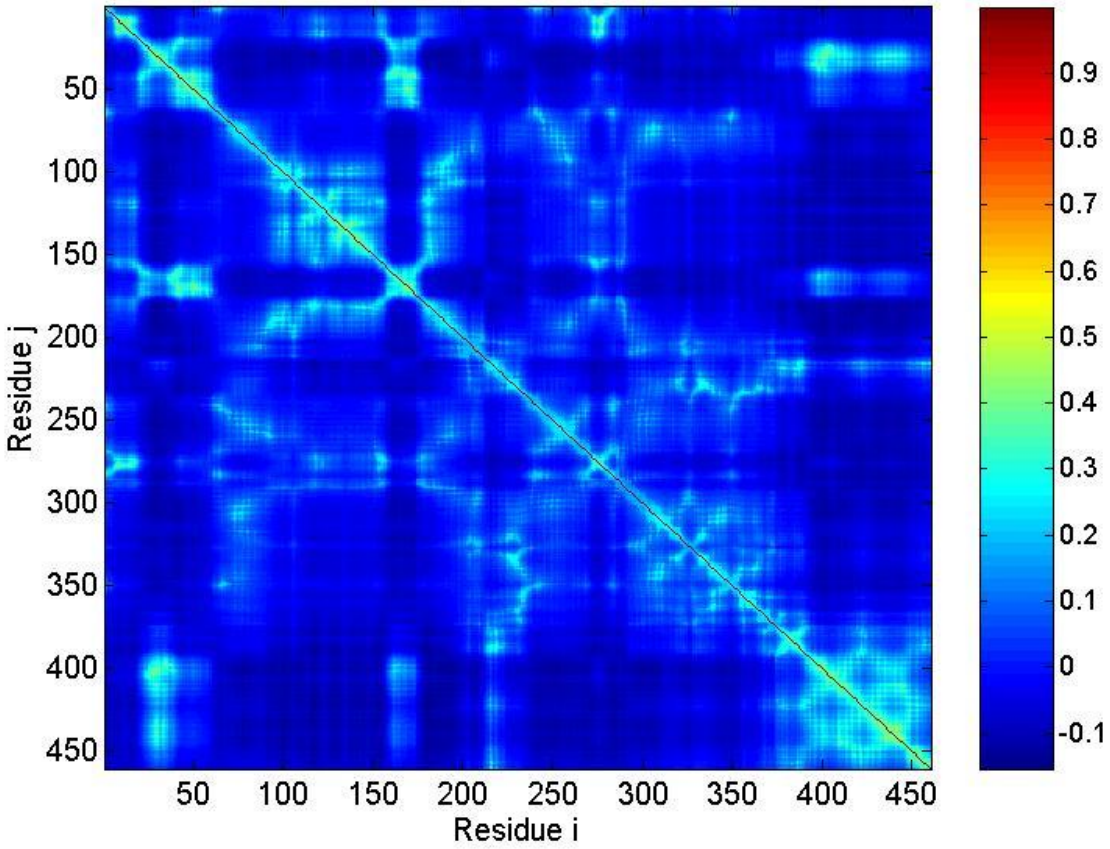
Isovitexin- Rhinovirus complex:

Elastic normal mode:

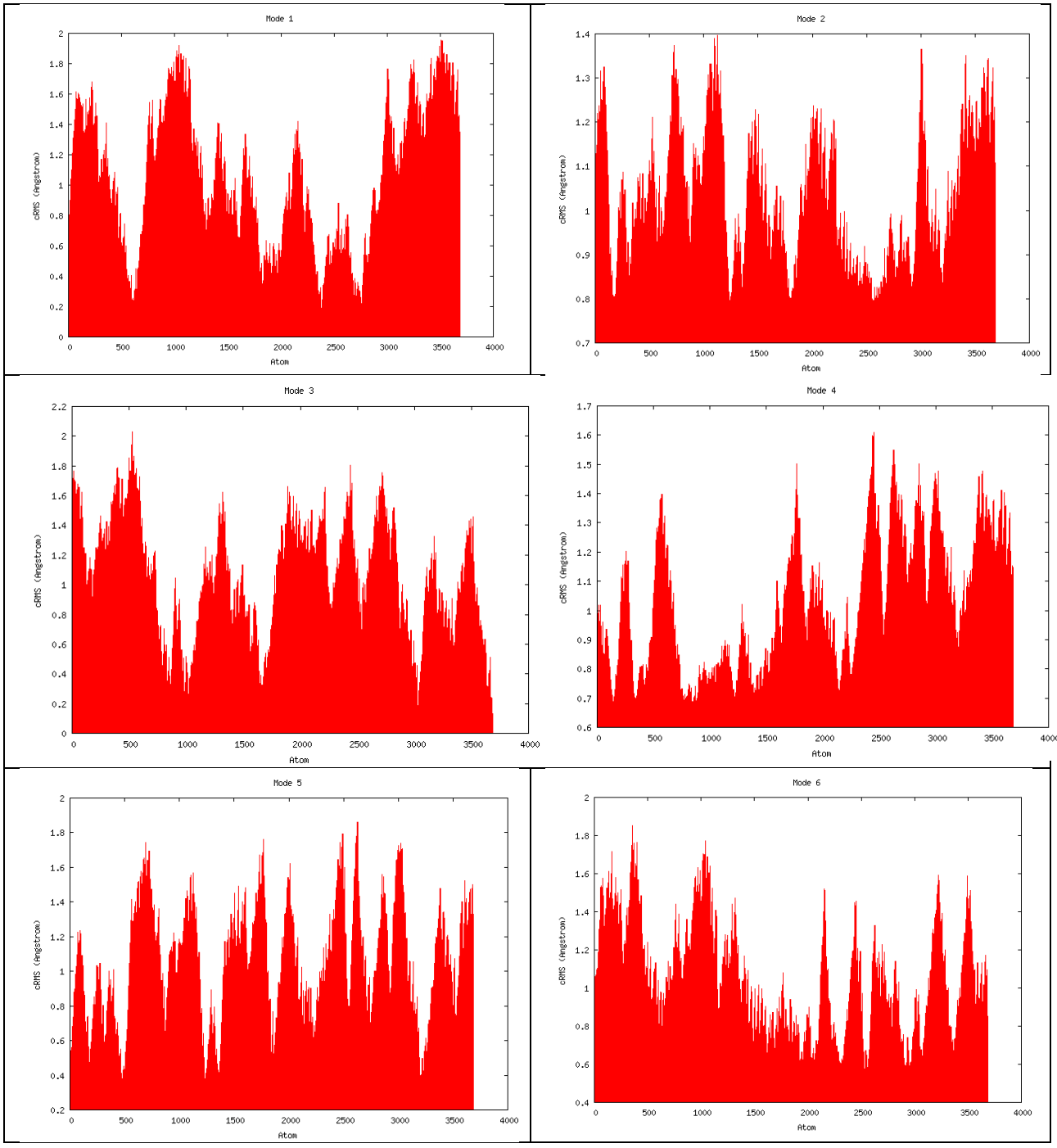


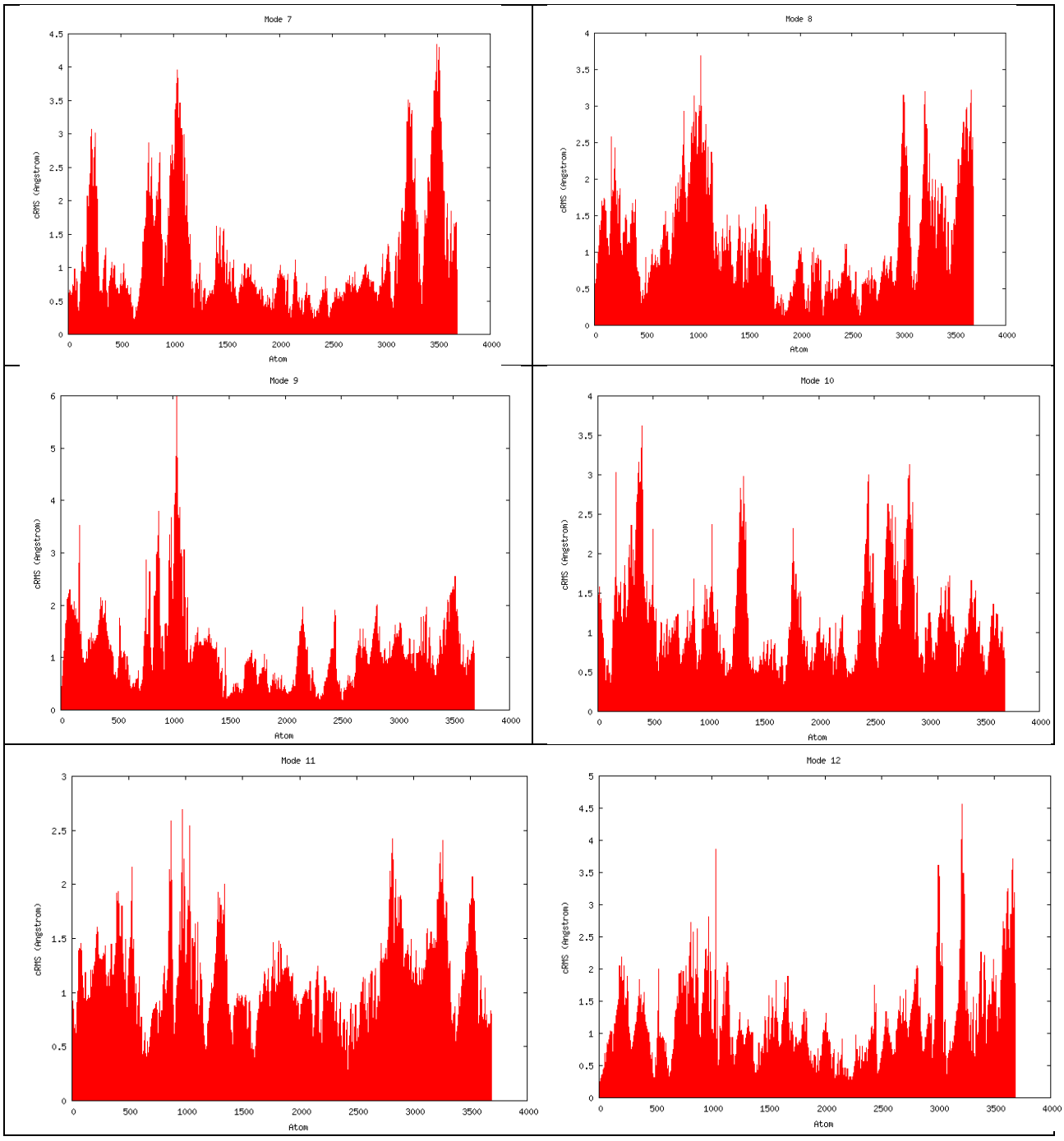


CC Plot:



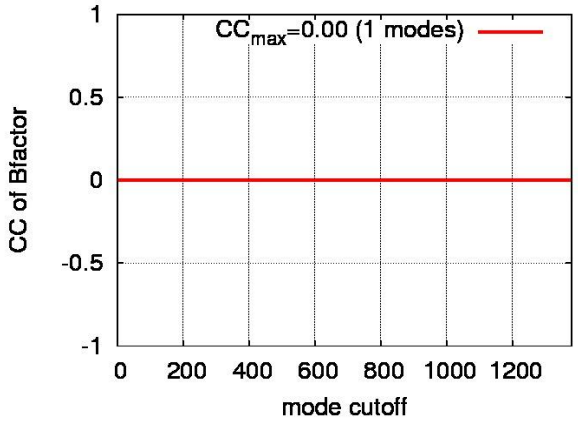
Normal mode analysis:

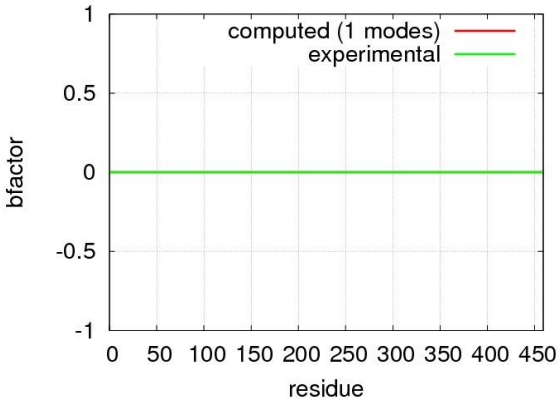



Mode 7
cRMS (Angstrom)
Atom
Mode 8
cRMS (Angstrom)
Atom
Mode 9
cRMS (Angstrom)
Atom
Mode 10
cRMS (Angstrom)
Atom
Mode 11
cRMS (Angstrom)
Atom
Mode 12
cRMS (Angstrom)
Atom

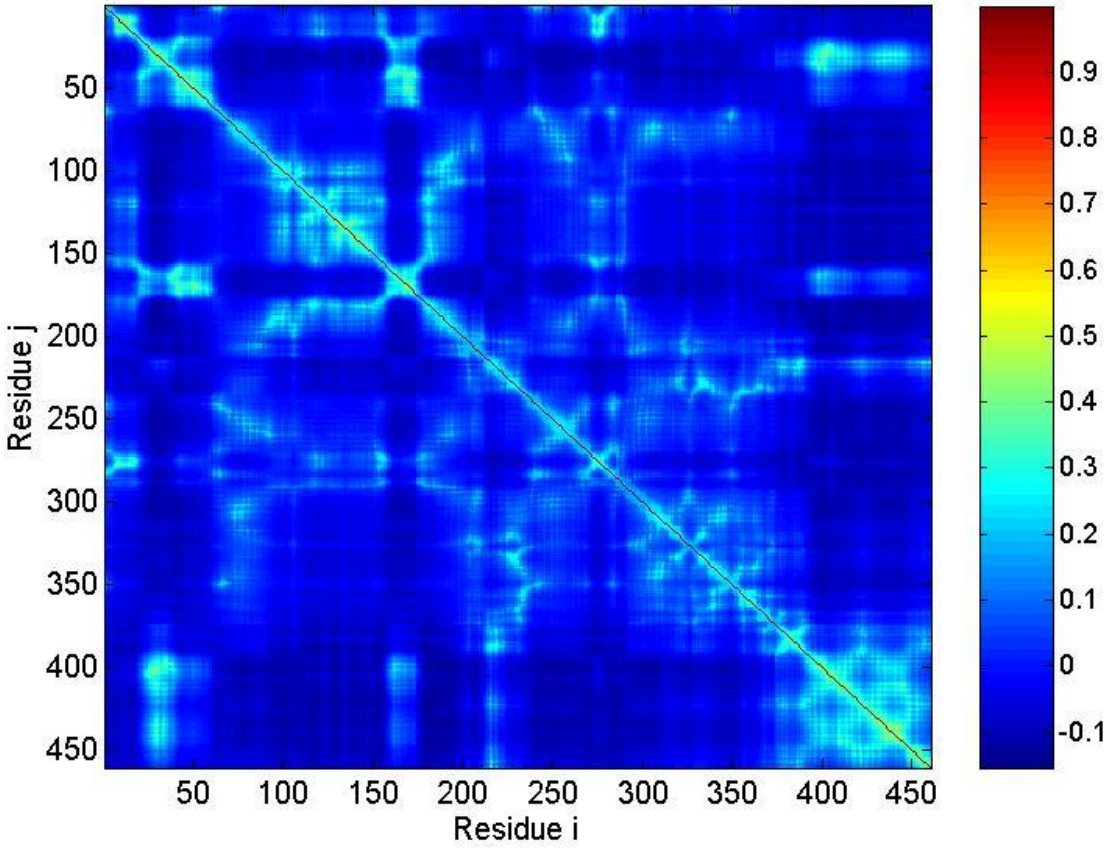
Alpha-amyrin Rhinovirus complex:

Elastic normal mode:

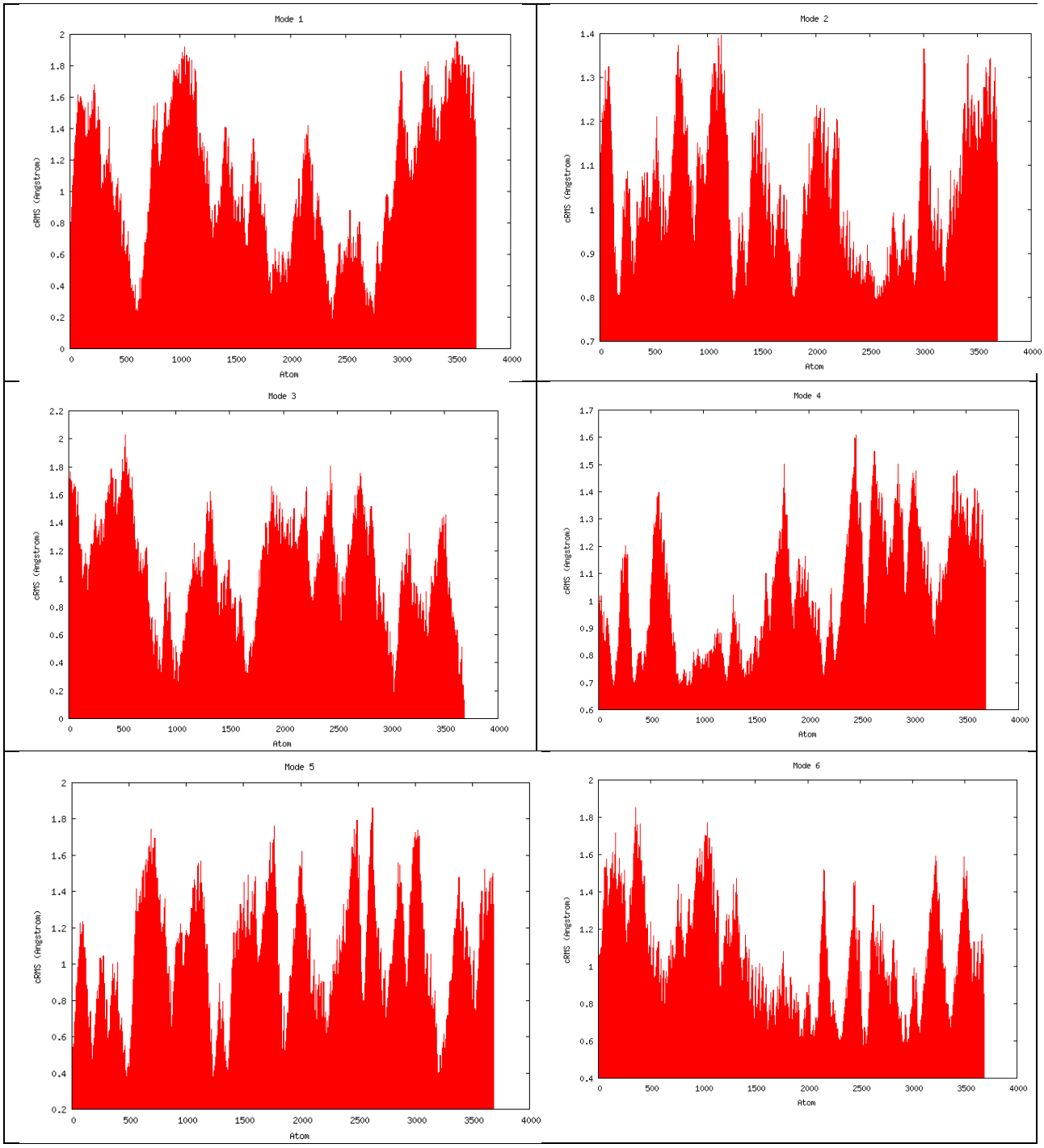


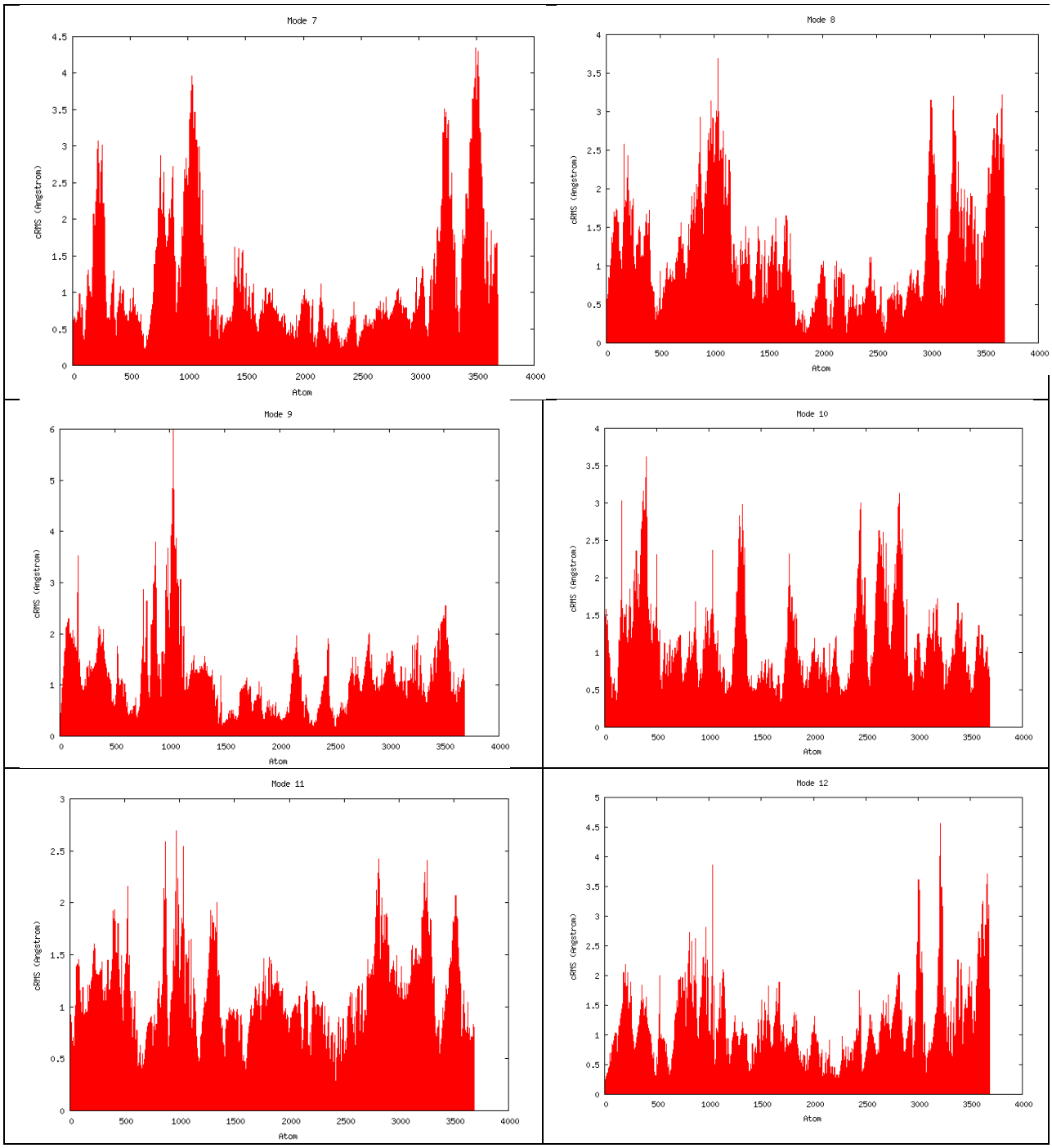


CC Plot



Normalmode analysis:





- **The structural dynamics models generated for the different protein-ligand complexes didn't show any remarkable difference.**

CHAPTER 5

CONCLUSION

5. CONCLUSION:

In this project, we successfully identified a molecule which could be used to manufacture a drug which might provide immunity against Rhinovirus caused common cold i.e., *Pinocembrin* obtained from honey. It can stop the RNA replication by the virus in the host by attaching to its Rdrp site.

Nonetheless, before manufacturing the drug, the molecule has to be further optimized. It also has to undergo *in vitro* and *in vivo* testing before being commercially manufactured. Apart from the chemical aspect of the viral infection, the physical aspect was also analyzed to make sure that the ligand didn't have any adverse physical effect on the body. As a result of which the structural dynamics of the Rhinovirus was successfully analyzed and compared to some of the structural proteins of the body and found to have similarity with collagen molecules. Hence, it has to be taken care of that the ligand doesn't affect the collagen molecules in the body anyway when administered. However, the dynamics of the protein was not remarkably affected by the binding of different ligands.

REFERENCES

1. Genetic clustering of all 102 human rhinovirus prototype strains: serotype 87 is close to human enterovirus 70: Carita Savolainen, Soile Blomqvist, Mick N. Mulders† and Tapani Hovi : Journal of General Virology (2002), 83, 333–340
2. Alignment of capsid protein VP1 sequences of all human rhinovirus prototype strains: conserved motifs and functional domains : Pia Laine, Soile Blomqvist, Carita Savolainen, Koen Andries and Tapani Hovi : Journal of General Virology (2006), 87, 129–138
3. New complete genome sequences of human rhinoviruses shed light on their phylogeny and genomic features : Caroline Tapparel, Thomas Junier, Daniel Gerlach, Samuel Cordey,
   Sandra Van Belle, Luc Perrin, Evgeny M Zdobnov and Laurent Kaiser : *BMC Genomics* 2007, **8**:224
4. Failure of Naturally Acquired Rhinovirus Infections to Produce Temporal Immunity to Heterologous Serotypes : T. E. MINOR, E. C. DICK, J. A. PETERSON, AND D. E. DOCHERTY : INFECTION AND IMMUNITY, Nov. 1974, P. 1192-1193
5. Co-ordinated Role of TLR3, RIG-I and MDA5 in the Innate Response to Rhinovirus in Bronchial Epithelium : Louise Slater, Nathan W. Bartlett, Jennifer J. Haas, Jie Zhu, Simon D. Message, Ross P. :Walton, Annemarie Sykes, Samer Dahdaleh, Deborah L. Clarke, Maria G. Belvisi, Onn M.Kon, Takashi Fujita, Peter K. Jeffery, Sebastian L. Johnston, Michael R. Edwards : PLoS Pathogens November 2010 | Volume 6 | Issue 11 | e1001178
6. Cross-Serotype Immunity Induced by Immunization with a Conserved Rhinovirus Capsid Protein: Nicholas Glanville., Gary R. Mclean., Bruno Guy, Valerie Lecouturier, Catherine Berry, Yves Girerd, Christophe Gregoire, Ross P. Walton, Rebecca M. Pearson, Tatiana Kebadze, Nicolas Burdin, Nathan W. Bartlett, Jeffrey W. Almond, Sebastian L. Johnston: PLOS Pathogens September 2013 | Volume 9 | Issue 9 | e1003669
7. Rhinovirus Uses a Phosphatidylinositol 4-Phosphate/Cholesterol Counter-Current for the Formation of Replication Compartments at the ER-Golgi Interface : Pascal S. Roulin, Mark Lötzerich, Federico Torta, Lukas B. Tanner, Frank J.M. van Kuppeveld, Markus R. Wenk, Urs F. Greber : Cell Host Microbe 2014 Nov 12;16(5):677-90
8. In-silico drug design: An approach which revolutionarised the drug discovery process A Wadood, N Ahmed , L Shah , A Ahmad , H Hassan , S Shams : OA Drug Design & Delivery 2013 Sep 01;1(1):3
9. In silico pharmacology for drug discovery: applications to targets and beyond, S Ekins, J Mestres and B Testa : British Journal of Pharmacology (2007) 152, 21–37
10. Human Rhinoviruses, Samantha E. Jacobs, Daryl M. Lamson, Kirsten St. George and Thomas J. Walsh : Clin. Microbiol. Rev. January 2013 vol. 26 no. 1 135-162

11. Integrating protein structural dynamics and evolutionary analysis with Bio3D: Lars Skjærven, Xin-Qiu Yao, Guido Scarabelli and Barry J Grant: BMC Bioinformatics 2014, 15:399
12. Functional aspects of protein flexibility, Kaare Teilum Æ Johan G. Olsen Æ Birthe B. Kragelund : Cell. Mol. Life Sci.DOI 10.1007/s00018-009-0014-6
13. Dynamic personalities of proteins, Katherine Henzler-Wildman1 & Dorothee Kern1: NATURE|Vol 450|13 December 2007|doi:10.1038/nature06522
14. Many Rhinovirus Serotypes Share the Same Cellular Receptor, GORDON ABRAHAMt AND RICHARD J. COLONNO : JOURNAL OF VIROLOGY, Aug. 1984, p. 340-345
15. Groupings of 90 rhinovirus serotypes. Crowell, R. L., and B. J. Landau. 1983.Infect. Immun. 37:642-647.
16. Receptors in the initiation of picornavirus infections, p. 1-42. In H. Fraenkel-Conrat and R. R. Wagner (ed.), Comprehensive virology, vol. Plenum Publishing Corp., New York.
17. Dick, E. C. 1968. Experimental infections of chimpanzees withhuman rhinovirus types 14 and 43. Proc. Soc. Exp. Biol. Med. 127:1079-1081.
18. Gauntt, C. J. 1974. Fragmentation of RNA in virus particles of rhinovirus type 14. J. Virol. 13:762-764.
19. Gauntt, C. J., M. M. Griffith, J. R. Sauck, R. H. Upson, and E. C. Carlson. 1975. Properties and origins of infectious rhinovirustype 14 particles of different buoyant densities. J. Virol.16:1265-1272.
20. Gerin, J. L., W. R. Richter, J. D. Fenters, and J. C. Holper.1968. Use of zonal ultracentrifuge systems for biophysical studies of rhinoviruses. J. Virol. 2:937-943.
21. Dodson G, Verma CS (2006) Protein flexibility: its role in structure and mechanism revealed by molecular simulations.Cell Mol Life Sci 63:207–219
22. Dodson GG, Lane DP, Verma CS (2008) Molecular simulations of protein dynamics: new windows on mechanisms in biology. EMBO Rep 9:144–150
23. Maragakis P, Lindorff-Larsen K, Eastwood MP, Dror RO, Klepeis JL, Arkin IT, Jensen MO, Xu H, Trbovic N, Friesner RA, Iii AG, Shaw DE (2008) Microsecond molecular dynamics simulation shows effect of slow loop dynamics on backbone amide order parameters of proteins. J Phys Chem B 112: 6155–6158
24. Linderstrøm-Lang K (1955) Deuterium exchange between peptides and water. Chem Soc Spec Publ 2:1–20
25. Hvidt A (1955) Deuterium exchange between ribonuclease and water. Biochim Biophys Acta 18:306–308
26. Teilum K, Olsen JG, Kragelund BB: Functional aspects of protein flexibility. Cell Mol Life Sci 2009, 66:2231–2247.
27. Henzler-Wildman K, Kern D: Dynamic personalities of proteins. Nature 2007, 450:964–972.

28. Henzler-Wildman KA, Thai V, Lei M, Ott M, Wolf-Watz M, Fenn T, Pozharski E, Wilson MA, Petsko GA, Karplus M, Hübner CG, Kern D: Intrinsic motions along an enzymatic reaction trajectory. Nature 2007, 450:838–844.
29. Vale RD, Milligan RA: The way things move: looking under the hood of molecular motor proteins. Science 2000, 288:88–95.
30. Yébenes H, Mesa P, Muñoz IG, Montoya G, Valpuesta JM: Chaperonins: two rings for folding. Trends Biochem Sci 2011, 36:424–432.